Norsk Regnesentral
NORWEGIAN COMPUTING CENTER

Note

Development of improved methods or basis for medical age assessments of minors and young adults

Final report

Note no SAMBA/14/16

Author Marion Haugen
Line Eikvil
Ingunn Fride Tvete
Sigrid Ingeborg Kvaal

Date April 28, 2016

Norwegian Computing Center

Norsk Regnesentral (Norwegian Computing Center, NR) is a private, independent, non-profit foundation established in 1952. NR carries out contract research and development projects in information and communication technology and applied statistical-mathematical modelling. The clients include a broad range of industrial, commercial and public service organisations in the national as well as the international market. Our scientific and technical capabilities are further developed in co-operation with The Research Council of Norway and key customers. The results of our projects may take the form of reports, software, prototypes, and short courses. A proof of the confidence and appreciation our clients have in us is given by the fact that most of our new contracts are signed with previous customers.

Norsk Regnesentral
Norwegian Computing Center
Postboks 114, Blindern
NO-0314 Oslo, Norway

Besøksadresse
Office address
Gaustadalléen 23
NO-0373 Oslo, Norway

Telefon · telephone
(+47) 22 85 25 00
Telefaks · telefax
(+47) 22 69 76 60

Internett · internet
www.nr.no
E-post · e-mail
nr@nr.no

Title	**Development of improved methods or basis for medical age assessments of minors and young adults**
Author	**Marion Haugen** <Marion.Haugen@nr.no> **Line Eikvil** <Line.Eikvil@nr.no> **Ingunn Fride Tvete** <Ingunn.Fride.Tvete@nr.no> **Sigrid Ingeborg Kvaal** <s.i.kvaal@odont.uio.no>

Date	April 28, 2016
Publication number	SAMBA/14/16

Abstract

This final report describes the results of the research conducted in the project «Development of improved methods or basis for medical age assessments of minors and young adults». The aim of the project has been to contribute to develop improved methods for measuring physical developments of a child or young adult and for assessing the chronological age based on such measurements, giving an improved basis for such assessments.

The project has brought together various international experts, both from the medical and odontology fields. The project has been an international collaboration with a group of partners with experience from both practical age estimation and research, from both of the currently used radiographic methods in Norway (wrist and teeth) and new MRI approaches. Several disciplines have been incorporated in the project: odontology, pediatrics, radiography, statistical modeling and image analysis.

Additional descriptions of methods and results from the statistical analyses and image analysis can be found in four additional notes.

The project is funded by The Norwegian Directorate of Immigration.

Keywords	Age assessment, wrist, teeth, manual, automated, agreement, differences, radiograph, MRI
Target group	Those officials who develop and use procedures for medical age assessment for determining the age of minors and young adults
Availability	Open
Project	FOUALDER
Project number	220642
Research field	Statistics for technology, industry and administration
Number of pages	62

Contents

List of Figures

List of Tables

1 Executive summary

This final report describes the results of the research conducted in the project «Development of improved methods or basis for medical age assessments of minors and young adults» (FOUALDER), funded by The Norwegian Directorate of Immigration (UDI). The aim of the project has been to contribute to develop improved methods for measuring physical developments of a child or young adult and for assessing the chronological age based on such measurements, giving an improved basis for such assessments.

The following summarizes the background, ethics, the project organization, the main activities and aims, the main results and conclusions, and recommendations.

1.1 Background

Unaccompanied minors[1] seeking asylum in Norway may be unaware of their exact date of birth, travel without appropriate identity documents or bring documents of uncertain provenance. With respect to rights and obligations for asylum seekers, there are significant legal differences between minors and adults. For Norwegian authorities, and in particular the UDI, it is important to determine whether the asylum seeker is older or younger than 18 years of age. A medical age assessment is carried out if this is impossible with reasonable certainty on the basis of other available credible information that is relevant for assessing the chronological age of the individual who claim to be an unaccompanied minor.

Medical age assessment in Norway is based on two biologically independent examinations (Gelbrich et al., 2015) performed at two separate institutions unaware of each other's conclusions. Skeletal maturity is evaluated from radiographs of the hand and wrist by the private radiology institute Unilabs (from August 1, 2013). Dental development is evaluated with a dental clinical examination and orthopantomograms (OPGs) of the jaw. Two different dentists at Institute of Clinical Dentistry, Faculty of Dentistry, University of Oslo perform the evaluation and a third dentist at the same institution assures the quality of the assessment. The two independent age assessments are summarized in a final age assessment at BarneSak AS (from August 1, 2013), which is sent to the UDI. Section 3.1.2 gives a more detailed description of the methods used for skeletal and dental age assesment. Figure 1 shows examples of a radiograph of the left hand and wrist, and an OPG of the jaw.

There is today no medical test or a group of tests that will absolutely and accurately let us know the exact chronological age of a human being. Due to biological variations exact age determination is with the present methods impossible, and only assessments can be made. There will always be biological variation and uncertainty associated with age assessments, and correctly expressing this uncertainty is just as important as the actual assessment. Age assessment methods which are both safer and more precise may be ben-

1. An unaccompanied minor (UAM) asylum seeker claims to be under the age of 18 years, and travel without parents or others with parental responsibility.

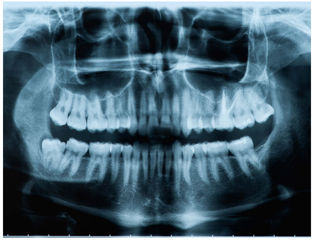
 

Figure 1. A radiograph of the left hand and wrist (on the left) and an OPG of the jaw (on the right). Photo: iStockphoto.com.

eficial in all situations where age should be determined (e.g., asylum seekers, criminal proceedings, sports competitions and identification) and for the society in general since it helps to protect minors who may otherwise inadvertently be lodged with adults or judged to be adults.

Immigration authorities in the European countries are not united on which methods to use for age assessment. There is a general agreement that the uncertainty due to biological variations may be reduced by using measurements from several physical properties (Gelbrich et al., 2015; Schmeling et al., 2008). In Norway and several other countries (e.g., Belgium and Germany), a combination of measurements from teeth and wrist is used. There are however different approaches for combining the results of these two medical age assessment methods to a final age assessment and how to use them with other types of relevant information.

Existing reference datasets for radiographs of the wrist and teeth do only contain one type of measurement per individual (only wrist or only teeth). Furthermore, they are based on limited datasets from few populations, they have age categories rather than continuous age, and the description of associated uncertainties is incomplete. Hence, there is a need for establishing sets of new reference data based on combinations of measurements (at least wrist and teeth) with better quality and understanding of the underlying data.

Updating of the currently available atlases using radiographs is however difficult, as a radiographic examination of healthy children with the sole aim of acquiring new standards is generally not permitted. New image atlases can be developed using nonionizing imaging techniques such as magnetic resonance imaging (MRI). Section 5 suggests an automatic approach to analyze the MR images shown in Figure 2.

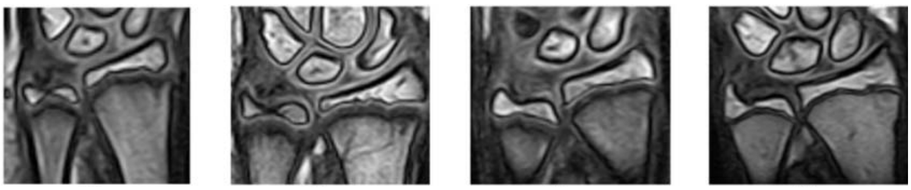


Figure 2. The development of the radial bone as observed in MRI-slices for four subsequent stages. Images from the University of Rome (Tomei et al., 2014).

Section 1.2 describes the ethics of the project, which has brought together various international experts, both from the medical and odontology fields. The project partners are given in section 1.3. The project has been organized into four main activities briefly described in section 1.4. Main results and conclusions are given in section 1.5. Finally, some recommendations to the UDI are given in section 1.6. Section 2 gives a summary in Norwegian. The main activities are described in more detail in the sections 3-6.

1.2 Ethics

Norwegian Computing Center (NR) has contacted the Norwegian Centre for Research Data (NSD) and sent an application for approval to the Regional Committees for Medical and Health Research Ethics (REK). Their conclusion was that there is no need to apply for permits to use the data in this project. Before NR has received any data for statistical analyses and image analysis, all the identities of the individuals have been removed.

1.3 Project organization

The project has been an international collaboration with a group of partners with experience from research and practical age assessment from both of the currently used radiographic methods in Norway (wrist and teeth), and new MRI approaches. The following disciplines have been incorporated in the project: odontology, pediatrics, radiography, statistical modeling and image analysis.

The project has consisted of the following partners:

- **Norwegian Computing Center (Norsk Regnesentral, NR):** Project manager
  - Marion Haugen, Senior Research Scientist
  - Ingunn Fride Tvete, Senior Research Scientist
  - Line Eikvil, Chief Research Scientist
  - André Teigland, Research Director for the SAMBA department and Deputy Director at NR
- **Institute of Clinical Dentistry, Faculty of Dentistry, University of Oslo**: Sigrid Ingeborg Kvaal, Associate Professor
- **Unilabs Norway**: Medical Doctor (MD) and Radiologist Tor Ole Kjellevand, Medical Manager
- **BarneSak AS**: MD and PhD Jens B. Grøgaard, Director at BarneSak AS and Senior Adviser at The Norwegian Directory of Health (former Head of Department of Pediatrics, Ullevaal University Hospital)
- **Institute of Dentistry, Queen Mary University of London**: Helen Liversidge, Reader in Dental Anthropology
- **Department of Radiology, University of Rome**: Professor Ernesto Tomei and MD Milvia Martino

All the groups working with the practical age assessments in Norway today have been included, i.e. Institute of Clinical Dentistry, Faculty of Dentistry, University of Oslo (dental age assessments), Unilabs Norway (skeletal age assessments) and BarneSak AS (final age assessments). This ensures that the front of the international research will be well known to the Norwegian experts in the field.

1.4 Main activities and aims

Here we give a brief description of the main activites and aims of the four work packages. These are all steps that are needed on the way to develop improved methods or basis for medical age assessments of minors and young adults. The work packages are described in more detail in the sections 3-6.

WP1 - Better understanding of combined data, comparison of skeletal and dental age assessment: The aim of this activity has been to get a better understanding of agreement and differences in the skeletal and dental age assessments performed in Norway, according to the asylum seeker's age being assessed to be above or below 18 years. The purpose has been to check for real differences and reveal any systematic differences between the two biologically independent age assessments over time, between genders and between citizenships. The analysis has not been a control of the accuracy of the methods used for age assessment, we have only compared and contrasted the assessment of maturity in the two biologically independent systems. We have analyzed the results from approximately 4 000 age assessments performed on unaccompanied minors in Norway from January 2010 to December 2014, separated in two data periods. This activity is part one of work package one (WP1) in the project and is described in section 3.

WP1 - Better understanding of combined data, comparison of manual and automated skeletal age assessment: We have tested BoneXpert, a software with continuous scale for automated estimation of skeletal age from a child's hand radiograph, on a sample of 100 left hand and wrist radiographs of boys. A manual skeletal age assessment of these boys has been performed earlier. The aim of this activity has been to understand agreement and differences between the two skeletal age assessments. This activity is part two of work package one (WP1) in the project and is described in section 3.

WP2 - Better understanding of dental variations between groups: The aim of this activity has been to identify and quantify differences in tooth development between ethnic groups and genders and how this impacts on estimated age from teeth. We have analyzed approximately 4 500 individuals with five different ethnicities. The individuals have known chronological age, in the age range 10-26 years, and developmental scores from third molar (wisdom tooth) stages and second molar stages. These are left mandibular teeth. We emphasize that the statistical analysis in WP2 separates from the dental age assessment performed on unaccompanied minor asylum seekers in Norway, as described in section 3.1.2. We have not considered how the dental age assessment of individuals is achieved in practice in this activity. WP2 provides a mathematical framework for estimating age distributions given molar stages for males and females for five ethnic groups. This activity is work package two (WP2) in the project and is described in section 4.

WP3 - Analysis of MRI-based data: At the Department of Radiology at the University of Rome, the research group of Professor Ernesto Tomei is working to develop an MRI-based approach for observing skeletal development, to overcome the problem of the need to collect new X-ray images to extend the current reference datasets. Within this activity we have investigated further the MRI-based methods suggested by Professor Tomei and his colleagues at the University of Rome. We have developed and explored methods based on automatic image analysis that can help in the process of analyzing MR images for bone age estimation. This activity is work package three (WP3) in the project and is described in section 5.

WP4 - Further development of new methods: Teeth and bones change shape/structure as they continue to grow until they reach full maturation in late adolescence or early adulthood. These changes are used to assess age. At present, the most commonly used methods for estimating age from teeth and bones assign a continuous process of maturation into discrete stages. WP4 has been a pilot study to explore methods that quantify dental maturation to improve reliability. The aim of this activity has been to establish a platform for further research in this field, by preparing and developing one or more PhD programs to collate these findings and establish a unified holistic approach to estimate age from teeth. This activity is work package four (WP4) in the project and is described in section 6.

1.5 Main results and conclusions

Here we give a brief description of the main results and conclusions from the four work packages.

WP1 - Better understanding of combined data, comparison of skeletal and dental age assessment: In section 3.1, agreement between age assessments based on skeletal maturity and age assessments based on dental development is defined in relation to the asylum seeker's age being assessed to be above or below 18 years. We consider it to be good agreement between the skeletal and dental age assessments performed on girls and boys in Norway from 2010 to 2014. The two age assessments are in agreement for approximately 80% of the girls in the two data periods, and for 86% and 78% of the boys. The skeletal age assessment is significantly higher than the dental age assessment in both data periods for girls, so there has been no change over time in the relationship between the two age assessments for girls. The skeletal age assessment is significantly higher than the dental age assessment in the first data period for boys. In the second data period, there is a larger proportion of boys with lower skeletal age assessment than dental age assessment. A shift in relationship between the two age assessments for boys may be explained by a younger age group in the second data period and different groups of unaccompanied minor asylum seekers in the two data periods.

WP1 - Better understanding of combined data, comparison of manual and automated skeletal age assessment: Our test of BoneXpert on a sample of 100 radiographs of boys (section 3.2) reveals an agreement above 90% between the manual and automated skeletal age assessment, in relation to the asylum seeker's age being assessed to be above or below 18 years. We consider this to be good agreement, which is a result of a precisely performed manual age assessment, a precise and reliable automated age assessment and good quality of the analyzed radiographs. However, there is a need to verify this finding on a larger dataset. A larger study is also necessary in order to decide an operational standard to determine whether the individual is above or below 18 years based on the continuous scale of BoneXpert. The error rates for the two outcomes of mismatches between the two skeletal age assessments can be balanced to acceptable levels by tuning a decision rule. Mismatches occur when the individual is 18 years or older from one of the skeletal age assessments and below 18 years from the other skeletal age assessment.

WP2 - Better understanding of dental variations between groups: We have provided a mathematical framework for estimating age distributions given molar stages and compared these across ethnic groups and genders. The analysis has been performed separately on the developmental scores from second and third molar stages for males and females from five different ethnic groups. The results in section 4 indicate differences across ethnic groups and also between genders, although weaker signs in the gender case. The results do not say anything about where the differences lie between the ethnic groups. More detailed ("head-to-head") comparisons between ethnic groups need to be done in order to determine which ethnic groups that differ.

WP3 - Analysis of MRI-based data: We have made a description of the manual method, and looked at studies done earlier on automated approaches for bone age estimation from X-rays and more recent studies dealing with this for MR images. On the basis of this, we have proposed a possible automatic approach, developed methods and implemented and tested this approach for bone age estimation of the radial bone on a set of MR images made at the University of Rome. The results reported in section 5, using the suggested automatic approach, are very promising. The conclusion so far is that for the current dataset, the automatic approach that we have suggested is able to give estimates with similar precision as that of the manually based method developed by the University of Rome. Hence, automation of the manual approach should be possible, and the suggested methods show promise. However, both the automatic method and the manual approach need to be validated on a larger independent dataset before very firm conclusions can be drawn. Such a large dataset does currently not exist, and an international effort is probably needed to collect such a set.

WP4 - Further development of new methods: We have done a literature review and initial exploration of alternative dental approaches for quantifying maturation and development. A project description for a PhD program that can look further into alternative dental approaches for quantifying maturation and development has been prepared. The literature review and a brief description of the PhD program is given in section 6.

1.6 Recommendations

The research in this project is a contribution to develop improved methods or basis for medical age assessments of minors and young adults. More work is needed to achieve the ultimate objective of age estimation methods which are both safer and more precise.

The research has resulted in the following recommendations:

- The development of hand bones and teeth are supposed to be biologically independent (Gelbrich et al., 2015) and are recommended used in age assessment procedures (Schmeling et al., 2008). We consider it to be good agreement between the age assessments based on skeletal maturity and the age assessments based on dental development performed on girls and boys in Norway from 2010 to 2014 (section 3.1). When there is a mismatch in the two age assessments, we recommend the UDI to put even more weight on other available age relevant information about the asylum seeker in the age assessment.
- Automated methods with continuous age can make the age assessment more objective and less dependent on the professionals performing the visual analysis. The results from the test of BoneXpert in section 3.2 are very promising. Prior to implementing BoneXpert in the age assessment process in Norway, this finding needs to be verified on a larger sample of radiographs. The continuous scale of BoneXpert provides more flexibility in the choice of a criterion to determine whether the individual is above or below 18 years. A larger study is needed to decide an operational standard for such a criterion. Due to biological variations, we will never get error rates for the two outcomes of mismatches between the two skeletal age assessments equal to zero. However, a decision rule can be tuned to balance the rates to acceptable levels.
- Use of MR images makes it possible to perform bone age estimation without the use of equipment generating ionizing radiation. This can again make it possible to acquire more up-to-date and larger reference datasets for future bone age estimation, which could be valuable. There is a need to verify both the automatic and the manual approach in section 5 on a larger independent MRI dataset. Prior to further development of these methods, it is necessary to collect a large dataset of MR images from healthy individuals of known age.
- To evaluate the operational consequences based on the findings from the project, we recommend the UDI to discuss this with the groups working with the practical age assessments in Norway today and those who will be working with this in the future.

1.7 References

Gelbrich B, Frerking C, Weiß S, Schwerdt S, et al. (2015). Combining wrist age and third molars in forensic age estimation: how to calculate the joint age estimate and its error rate in age diagnostics. Ann Hum Biol., Vol. 42(4):389-396.

Schmeling A, Grundmann C, Fuhrmann A, Kaatsch HJ, et al. (2008). Criteria for age estimation in living individuals. Int J Legal Med., Vol. 122(6):457-460.

Tomei E, Semelka RC, Nissman D. (2014). Text-Atlas of Skeletal Age Determination: MRI of the Hand and Wrist in Children. ISBN: 978-1-118-69227-1. Wiley-Blackwell.

2 Sammendrag

Denne sluttrapporten beskriver resultatene av forskningen i prosjektet «Utvikling av forbedrede metoder og grunnlag for medisinske aldersvurderinger av mindreårige og unge voksne» (FOUALDER), finansiert av Utlendingsdirektoratet (UDI). Målet med prosjektet har vært å bidra til å utvikle forbedrede metoder for måling av skjelettmodning og tannutvikling hos ungdom og unge voksne. Den kronologiske alderen til et individ med ukjent alder vurderes i Norge basert på skjelettmodning og tannutvikling, og forbedrede metoder for å måle disse vil derfor kunne gi et bedre grunnlag for å vurdere denne alderen.

I det følgende oppsummeres bakgrunn, etikk, prosjektorganisering, hovedaktiviteter og mål, de viktigste resultater og konklusjoner, og anbefalinger.

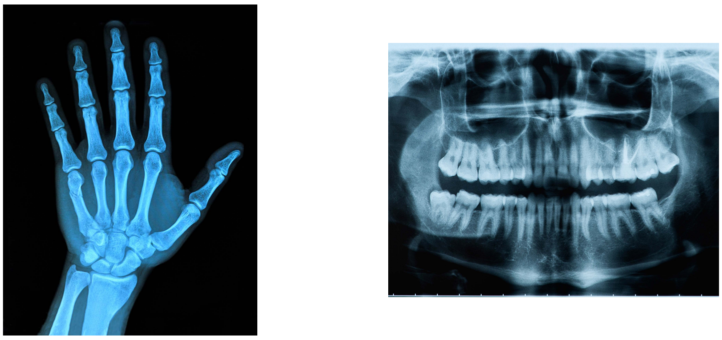
2.1 Bakgrunn

Mange enslige mindreårige asylsøkere (EMA)[2] kjenner ikke til egen alder, eller har manglende eller usikre ID-papirer. En søkers alder er av betydning med tanke på både rettigheter og plikter i samfunnet. For norske myndigheter og spesielt UDI er det viktig å fastslå om asylsøkeren er over eller under 18 år. En medisinsk aldersvurdering benyttes dersom det ikke er mulig å fastsette med rimelig sikkerhet hvor gammel en EMA er, på grunnlag av annen tilgjengelig troverdig informasjon som er relevant for å vurdere den kronologiske alderen til individet som hevder å være en enslig mindreårig.

Den medisinske aldersvurderingen i Norge består av to biologisk uavhengige undersøkelser (Gelbrich et al., 2015) som utføres på to adskilte institusjoner, som ikke kjenner til hva den andre gjør eller konkluderer med. Skjelettmodningen vurderes med en røntgenundersøkelse av hånd og håndrot av det private røntgeninstituttet Unilabs Norge (fra 1. august 2013). Den odontologiske aldersvurderingen omfatter en klinisk undersøkelse av tenner og munnhule, og det tas et stort røntgenbilde som inkluderer alle tennene, tannrøttene og kjevene, et såkalt ortopantomogram (OPG) eller panoramarøntgen. Institutt for klinisk odontologi, Det odontologiske fakultet, Universitetet i Oslo er ansvarlig for aldersvurderinger basert på tannutvikling. På bakgrunn av undersøkelsen så vurderer to ulike tannleger asylsøkerens alder og en tredje tannlege kvalitetssikrer resultatet. BarneSak AS har fra 1. august 2013 sammenstilt resultatene av disse to uavhengige undersøkelsene til en endelig medisinsk aldersvurdering, som sendes til UDI. Metodene som brukes i skjelett- og tannvurderingen er beskrevet i avsnitt 3.1.2. Figur 3 viser eksempler på et røntgenbilde av hånd og håndrot, og et OPG av alle tennene, tannrøttene og kjevene.

Det finnes ingen medisinsk test eller gruppe av tester som helt nøyaktig kan fastsette den kronologiske alderen til et individ. Nøyaktig aldersbestemmelse er med de eksisterende metoder umulig på grunn av biologiske variasjoner, og det er kun aldersvur-

2. Enslig mindreårig asylsøker (EMA) er en asylsøker som oppgir å være under 18 år, og som ikke har følge av foreldre eller andre med foreldreansvar.



Figur 3. Et røntgenbilde av venstre hånd og håndrot (til venstre), og et OPG som viser alle tennene, tannrøttene og kjevene (til høyre). Foto: iStockphoto.com.

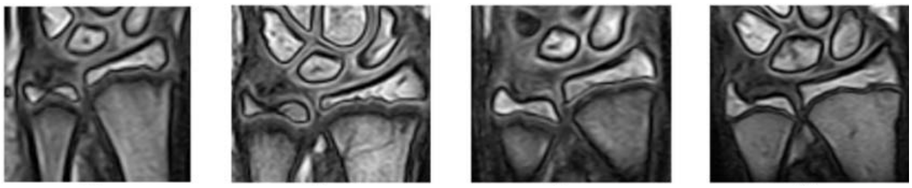
deringer som kan bli utført. Det vil alltid være biologisk variasjon og usikkerhet knyttet til aldersvurderingen og en riktig beskrivelse av denne usikkerheten er like viktig som selve vurderingen. Mer presise metoder for aldersvurdering er nyttig i alle situasjoner der alder skal fastsettes (f.eks. asylsaker, straffesaker, idrettskonkurranser og identifisering) og for samfunnet generelt. Det bidrar til å beskytte mindreårige som ellers utilsiktet kan bli innkvartert sammen med voksne asylsøkere eller bli behandlet som voksne.

Immigrasjonsmyndighetene i de europeiske landene er ikke samlet om hvilke metoder som skal brukes for aldersvurdering, men for å redusere usikkerheten knyttet til biologisk variasjon så er det en generell enighet om at målinger fra flere enn én fysisk egenskap skal brukes ved slike vurderinger (Gelbrich et al., 2015; Schmeling et al., 2008). I Norge og flere andre land (f.eks. Belgia og Tyskland) brukes en kombinasjon av målinger fra tenner og håndledd. Det finnes imidlertid forskjellige måter å kombinere resultatene fra de to medisinske metodene for aldersvurdering til en endelig aldersvurdering, og ulike tilnærminger til hvordan de brukes sammen med annen relevant informasjon.

Eksisterende referanse datasett for røntgenbilder av håndledd og tenner inneholder bare én type måling per individ (bare håndledd eller bare tenner). De er basert på begrensede datasett fra noen få populasjoner, de har alderskategorier i stedet for kontinuerlig alder, og beskrivelsen av den tilhørende usikkerheten er ufullstendig. Det er derfor behov for å etablere et sett av nye referansedata basert på kombinasjoner av målinger (i det minste håndledd og tenner), med bedre kvalitet og forståelse av de underliggende data.

Oppdatering av eksisterende bildeatlas som bruker røntgenbilder er vanskelig siden det generelt ikke er tillatt å foreta en røntgenundersøkelse av friske barn til ikke-medisinske formål. Nye bildeatlas kan utvikles ved hjelp av ikke-ioniserende teknikker slik som MR. I avsnitt 5 har vi sett på om og hvordan det kan utvikles en metode for automatisk estimering av skjelettalder fra MR-bilder vist i figur 4.

Etikken i prosjektet er beskrevet i avsnitt 2.2. Prosjektet har samlet ulike internasjonale eksperter, både fra medisin og odontologi. Prosjektpartnerne er gitt i avsnitt 2.3. Prosjektet har vært organisert i fire hovedaktiviteter som er kort beskrevet i avsnitt 2.4. Hovedresultater og konklusjoner er gitt i avsnitt 2.5. Noen anbefalinger til UDI er gitt i avsnitt 2.6. Hovedaktivitetene er beskrevet i mer detalj i avsnittene 3-6.



Figur 4. Fire påfølgende utviklingstrinn av radialbeinet. Bilder fra Tomei et al. (2014).

2.2 Etikk

Norsk Regnesentral (NR) har vært i kontakt med Norsk senter for forskningsdata (NSD) og sendt søknad om forhåndsgodkjenning til Regionale komiteer for medisinsk og helsefaglig forskningsetikk (REK). NSD og REK har uttalt at det ikke er behov for å søke om tillatelser til å bruke dataene i dette prosjektet og det kreves ingen prosjektgodkjenning fra REK for å gjennomføre prosjektet. Alle data som NR har mottatt til statistisk analyse og bildeanalyse har vært avidentifisert.

2.3 Prosjektorganisering

Prosjektet har vært et internasjonalt samarbeid med en gruppe av partnere med erfaring fra forskning og praktiske aldersvurderinger, både fra røntgenmetodene som brukes i Norge i dag (skjelett og tenner) og nye MR-metoder. Følgende faggrupper har vært med i prosjektet: odontologi, pediatri, radiografi, statistisk modellering og bildeanalyse.

Prosjektet har bestått av følgende partnere:

- **Norsk Regnesentral (NR):** Prosjektleder
  - Marion Haugen, Seniorforsker
  - Ingunn Fride Tvete, Seniorforsker
  - Line Eikvil, Sjefsforsker
  - André Teigland, Forskningssjef for SAMBA avdelingen og Assisterende direktør ved NR
- **Institutt for klinisk odontologi, Det odontologiske fakultet, Universitetet i Oslo:** Sigrid Ingeborg Kvaal, Førsteamanuensis
- **Unilabs Norge:** Lege og radiolog Tor Ole Kjellevand, Medisinsk sjef
- **BarneSak AS**: Lege og dr.med. Jens B. Grøgaard, Daglig leder i BarneSak AS og Seniorrådgiver på Helsedirektoratet (tidligere Klinikksjef på barneklinikken Ullevaal)
- **Institutt for odontologi, Queen Mary University of London:** Helen Liversidge, Førsteamanuensis i dentalantropologi
- **Avdeling for radiologi, Universitetet i Roma:** Professor Ernesto Tomei og lege Milvia Martino

Alle aktørene som jobber med de praktiske aldersvurderingene i Norge i dag har vært med i prosjektet, dvs. Institutt for klinisk odontologi, Det odontologiske fakultet, Universitetet i Oslo (tannvurderinger), Unilabs Norge (skjelettvurderinger) og BarneSak AS (endelige aldersvurderinger). Dette gjør at resultatene av forskningen i prosjektet vil være kjent for de norske ekspertene på området.

2.4 Hovedaktiviteter og mål

Her gir vi en kort beskrivelse av de viktigste aktivitetene og målene for de fire arbeidspakkene. Alle aktivitetene er nødvendige steg for å utvikle forbedrede metoder og grunnlag for medisinske aldersvurderinger av ungdom og unge voksne. Arbeidspakkene er beskrevet i mer detalj i avsnittene 3-6.

Arbeidspakke 1 - Bedre forståelse av kombinerte data, sammenligning av skjelett- og tannvurdering: Målet med dette arbeidet har vært å se på samsvar og avvik mellom aldersvurderinger basert på skjelettmodning og aldersvurderinger basert på tannutvikling, i henhold til om asylsøkers alder vurderes til å være over eller under 18 år. Hensikten har vært å avdekke eventuelle systematiske forskjeller mellom de to biologisk uavhengige aldersvurderingene, både over tid, mellom kjønn og mellom opprinnelsesland. Analysen har ikke vært en kontroll av treffsikkerheten til metodene som brukes til aldersvurdering, kun av hvor godt de samsvarer med hverandre. Vi har analysert resultatene fra ca. 4.000 aldersvurderinger utført på enslige mindreårige asylsøkere i Norge fra januar 2010 til desember 2014, inndelt i to dataperioder. Dette arbeidet er beskrevet i avsnitt 3.

Arbeidspakke 1 - Bedre forståelse av kombinerte data, sammenligning av manuell og automatisert skjelettvurdering: Vi har testet BoneXpert, en programvare med kontinuerlig skala for automatisert estimering av skjelettalder fra røntgenbilder, på 100 utvalgte røntgenbilder av gutter. En manuell skjelettvurdering av disse guttene har blitt utført tidligere. Målet med dette arbeidet har vært å se på samsvar og avvik mellom manuell og automatisert skjelettvurdering. Dette arbeidet er beskrevet i avsnitt 3.

Arbeidspakke 2 - Bedre forståelse av tannvariasjon mellom grupper: Målet med dette arbeidet har vært å identifisere og kvantifisere forskjeller i tannutvikling mellom etniske grupper og kjønn, og hvilken betydning dette har for aldersestimering basert på tenner. Vi har analysert et datamateriale på ca. 4.500 individer med fem ulike etnisiteter. Individene har kjent kronologisk alder (fra 10 år til 26 år) og fastsatt utviklingsstadie fra visdomstannen (tredje molar) og 12-års jekselen (andre molar) på venstre side i underkjeven. Vi understreker at den statistiske analysen i arbeidspakke 2 skiller seg fra den praktiske tannvurderingen som utføres på enslige mindreårige asylsøkere i Norge i dag, som beskrevet i avsnitt 3.1.2. Vi har ikke sett på hvordan aldersvurderinger basert på tannutvikling gjøres i praksis i denne aktiviteten, men anvendt et matematisk rammeverk for å estimere aldersfordelinger gitt utviklingsstadier fra visdomstannen og 12-års jekselen på venstre side i underkjeven, både for menn og kvinner for fem ulike etniske grupper. Dette arbeidet er beskrevet i avsnitt 4.

Arbeidspakke 3 - Analyse av MR-baserte data: På Avdeling for radiologi ved Universitetet i Roma arbeider forskningsgruppen til professor Ernesto Tomei med å utvikle en MR-basert tilnærming for å observere skjelettutvikling. Dette gjøres for å unngå problemet med å måtte samle inn nye røntgenbilder for oppbygging av nye og større bildeatlas av skjelettutvikling. Målet med dette arbeidet har vært å se på om og hvordan det kan utvikles en metode for automatisk estimering av skjelettalder fra MR-bilder med utgangspunkt i den manuelt baserte metoden utviklet ved Universitetet i Roma. Dette arbeidet er beskrevet i avsnitt 5.

Arbeidspakke 4 - Alternative metoder til odontologisk aldersvurdering ved bruk av 3D-bilder: Når tenner og skjelett vokser så endres form og struktur. Endringen foregår inntil tenner og skjelett når full modning i slutten av tenårene eller tidlig voksen alder. Aldersvurdering skjer ved fastsatte metoder som kobler utviklingsnivået til skjelett og tenner til en viss alder via diskrete tilstander. Dette arbeidet har vært en pilotstudie for å utforske metoder som kan forbedre påliteligheten til tannvurderinger. Målet med denne aktiviteten har vært å etablere en plattform for videre forskning på dette feltet ved å forberede og utvikle ett eller flere PhD-programmer for å sammenstille disse funnene, og etablere en helhetlig tilnærming for å bruke tenner til aldersvurdering. Arbeidet er beskrevet i avsnitt 6.

2.5 Hovedresultater og konklusjoner

Her gir vi en kort beskrivelse av de viktigste resultatene og konklusjonene fra de fire arbeidspakkene.

Arbeidspakke 1 - Bedre forståelse av kombinerte data, sammenligning av skjelett- og tannvurdering: I avsnitt 3.1 er samsvar mellom aldersvurderinger basert på skjelettmodning og aldersvurderinger basert på tannutvikling definert i henhold til om asylsøkers alder vurderes til å være over eller under 18 år. Vi mener at det er godt samsvar mellom skjelett- og tannvurdering utført på jenter og gutter i Norge fra 2010 til 2014. Samsvaret for jenter er ca. 80% i begge dataperioder og for gutter er det hhv. 86% og 78% i de to dataperiodene. Skjelettvurderingen er signifikant høyere enn tannvurderingen i begge dataperioder for jenter så det har ikke skjedd en endring over tid i forholdet mellom de to aldersvurderingene for jenter. I den første dataperioden er skjelettvurderingen signifikant høyere enn tannvurderingen for gutter. I den andre dataperioden er det en større prosentandel gutter som har fått lavere skjelettvurdering enn tannvurdering. En yngre aldersgruppe i den andre dataperioden og ulike grupper av enslige mindreårige asylsøkere i de to dataperiodene kan være med å forklare et skifte i forholdet mellom de to aldersvurderingene for gutter.

Arbeidspakke 1 - Bedre forståelse av kombinerte data, sammenligning av manuell og automatisert skjelettvurdering: Testen vi har utført av BoneXpert på 100 utvalgte røntgenbilder viser at det er over 90% samsvar mellom manuell og automatisert skjelettvurdering, i henhold til om asylsøkers alder vurderes til å være over eller under 18 år (avsnitt 3.2). Vi mener at dette er godt samsvar, som skyldes at den manuelle skjelettvurderingen er presist utført, at BoneXpert er presis og gir en pålitelig skjelettvurdering, og at det er god kvalitet på de analyserte røntgenbildene. Det er behov for en større studie for å verifisere det gode samsvaret og for å bestemme en operativ standard, dvs. en grense, for å avgjøre om individet er over eller under 18 år fra den kontinuerlige skalaen til BoneXpert. Med et større utvalg røntgenbilder av individer med kjent kronologisk alder kan feilratene for de to utfallene der det ikke er samsvar mellom de to skjelettvurderingene beregnes for ulike grenser. Individet er 18 år eller eldre fra den ene skjelettvurderingen og under 18 år fra den andre skjelettvurderingen når det ikke er samsvar. Ved å veie ulike hensyn kvantitativt opp mot hverandre kan feilratene i begge retninger balanseres til akseptable nivåer.

Arbeidspakke 2 - Bedre forståelse av tannvariasjon mellom grupper: Vi har anvendt et matematisk rammeverk for å estimere aldersfordelinger gitt utviklingsstadier fra visdomstannen (tredje molar) og 12-års jekselen (andre molar) på venstre side i underkjeven. For å kunne sammenligne aldersfordelingene over etniske grupper og mellom kjønn, har vi gjort separate analyser for menn og kvinner med fem ulike etnisiteter. Resultatene i avsnitt 4 indikerer en forskjell mellom etniske grupper og mellom kjønn, men forskjellene mellom kjønn er dog mindre enn mellom etniske grupper. Analysen vi har gjort sier ingenting om hvor denne forskjellen mellom etniske grupper ligger. Mer detaljerte sammenligninger (én etnisk gruppe mot en annen) må gjøres for å se på hvilken gruppe eller hvilke grupper som skiller seg ut.

Arbeidspakke 3 - Analyse av MR-baserte data: Vi har satt oss inn i hvordan den manuelle metoden virker og laget en beskrivelse av denne. Videre har vi sett på studier som er gjort tidligere på automatisert beregning av skjelettalder fra røntgenbilder, samt helt ferske studier som tar for seg dette for MR-bilder. På bakgrunn av dette har vi foreslått en mulig automatisk fremgangsmåte, utviklet metoder og implementert og testet ut dette for beregning av skjelettalder fra radialbeinet på et sett med bilder fra MR-opptak gjort ved Universitetet i Roma. Resultatene i avsnitt 5 viser at den foreslåtte automatiske metoden, for det datasettet vi har sett på, er i stand til å gi estimater innenfor samme feilmargin som den manuelt baserte metoden. Metoden virker lovende og bør ha potensiale til å brukes i estimering av skjelettalder. Datasettet vi har hatt tilgjengelig har imidelrtid vært begrenset. Både den automatiske tilnærmingen og den manuelle MR-metoden må valideres på et større og uavhengig datasett. Et slikt stort MR-datasett finnes ikke per i dag og en samlet internasjonal innsats kan være viktig for å etablere et slikt sett.

Arbeidspakke 4 - Alternative metoder til odontologisk aldersvurdering ved bruk av 3D-bilder: Vi har gjort en litteraturgjennomgang og innledende utforskning av alternative tilnærminger for å kvantifisere modning og utvikling av tenner. En prosjektbeskrivelse for et PhD-program som kan se mer på alternative tilnærminger for å kvantifisere modning og utvikling av tenner er utarbeidet. Arbeidspakken er beskrevet i avsnitt 6.

2.6 Anbefalinger

Forskningen i dette prosjektet er et bidrag til å utvikle forbedrede metoder og grunnlag for medisinske aldersvurderinger av ungdom og unge voksne. Mer arbeid er nødvendig for å oppnå hovedmålet om mer presise metoder for aldersvurdering.

Forskningen har resultert i følgende anbefalinger:

- Utviklingen av skjelettet i hånden og tenner er biologisk uavhengige (Gelbrich et al., 2015) og er anbefalt brukt til aldersvurderinger (Schmeling et al., 2008). Vi mener at det er godt samsvar mellom aldersvurderinger basert på skjelettmodning og aldersvurderinger basert på tannutvikling utført på jenter og gutter i Norge fra 2010 til 2014 (avsnitt 3.1). Når disse to aldersvurderingene ikke er i samsvar, er anbefalingen til UDI å legge ekstra vekt på øvrig aldersrelevant informasjon om asylsøkeren i aldersvurderingen.

- Automatiserte metoder med kontinuerlig alder kan gjøre aldersvurderingen mer objektiv og reproduserbar, og mindre avhengig av ekspertene som utfører de visuelle analysene. Resultatene i avsnitt 3.2 fra en test av BoneXpert er lovende. Før BoneXpert eventuelt skal brukes til aldersvurderinger i Norge må dette funnet bekreftes på et større utvalg røntgenbilder. BoneXpert har en kontinuerlig skala og dette gir mer fleksibilitet i valg av kriterium for å avgjøre om individet er over eller under 18 år. En større studie er nødvendig for å bestemme en operativ standard for et slikt kriterium. Uansett hvor grensen settes vil det, grunnet biologisk variasjon, være andeler feilklassifikasjon for de to utfallene som ikke er i samsvar. Ved å veie ulike hensyn kvantitativt opp mot hverandre kan feilratene i begge retninger balanseres til akseptable nivåer.

- Med MR kan en skjelettvurdering utføres uten bruk av ioniserende stråling for å fremstille bildene. Dette gjør det mulig å kunne bygge opp nye og større bildeatlas av skjelettutvikling, som kan være verdifullt for fremtidig aldersestimering fra skjelett. Både den automatiske og den manuelle tilnærmingen i avsnitt 5 må valideres på et større og uavhengig datasett. Før videre utvikling av disse metodene, er det nødvendig å samle inn et stort datasett av MR-bilder fra friske individer med kjent alder.

- For å vurdere de operative konsekvensene basert på resultatene fra prosjektet, anbefaler vi UDI å diskutere dette med aktørene som jobber med de praktiske aldersvurderingene i Norge i dag og de som skal jobbe med dette i fremtiden.

2.7 Referanser

Gelbrich B, Frerking C, Weiß S, Schwerdt S, et al. (2015). Combining wrist age and third molars in forensic age estimation: how to calculate the joint age estimate and its error rate in age diagnostics. Ann Hum Biol., Vol. 42(4):389-396.

Schmeling A, Grundmann C, Fuhrmann A, Kaatsch HJ, et al. (2008). Criteria for age estimation in living individuals. Int J Legal Med., Vol. 122(6):457-460.

Tomei E, Semelka RC, Nissman D. (2014). Text-Atlas of Skeletal Age Determination: MRI of the Hand and Wrist in Children. ISBN: 978-1-118-69227-1. Wiley-Blackwell.

3 WP1: Better understanding of combined data

The objective of this activity has been to analyze existing data on the results from all age assessments performed on unaccompanied minor asylum seekers in Norway from January 2010 to December 2014 (approximately 4 000 age assessments), to understand agreement and differences between age assessments based on skeletal maturity and dental development over time, between genders, between citizenships and between methods. We have analyzed three different methods for age assessment and the activity has therefore been separted in two parts:

(i) Analysis of agreement and differences between age assessment based on radiographs of the left hand and wrist, analyzed with the method of Greulich and Pyle (1959), and age assessment based on a dental clinical examination and orthopantomograms (OPG) of the jaw, analyzed with the methods of Haavikko (1970); Kvaal et al. (1995); Liversidge (2008). The analysis is described in section 3.1.

(ii) Analysis of agreement and differences between a manual and an automated approach for skeletal age assessment. A sample of 100 left hand and wrist radiographs, previously analyzed with the method of Greulich and Pyle (1959), has been reanalyzed using BoneXpert, a software for automated measurement of bone age from a child's hand and wrist radiograph. The analysis is described in section 3.2.

> The following summarizes the work that has been done within WP1. More results can be found in two different notes (written in Norwegian): «Sammenligning av to metoder for aldersvurdering av enslige mindreårige asylsøkere» (Haugen et al., 2016) and «Sammenligning av manuell og automatisert skjelettvurdering» (Haugen, Grøgaard and Kjellevand, 2016).

Sigrid Ingeborg Kvaal, Tor Ole Kjellevand and Jens B. Grøgaard have contributed considerably to this work package with their expert knowledge on the field. They have provided the data, participated in project meetings and revised the notes critically. Tor Ole Kjellevand has performed the test of BoneXpert. Marion Haugen has conducted the statistical analysis and written the notes.

3.1 Comparison of skeletal and dental age assessment

3.1.1 Background

There are two data periods (January 2010 - July 2013 and August 2013 - December 2014), depending on the institution performing the skeletal age assessment. Until July 31, 2013, the skeletal age assessments were performed at Pediatric Radiology, Ullevaal University Hospital, and is now performed at the private radiology institute Unilabs Norway (from August 1, 2013). The dental age assessments are performed at Institute of Clinical Dentistry, Faculty of Dentistry, University of Oslo. Department of Pediatrics, Oslo University Hospital, Ullevaal, compiled the results of the two independent examinations into a final age assessment from January 1, 2010 until July 31, 2013. The final age assessment is now performed at BarneSak AS (from August 1, 2013).

The skeletal age assessments have been performed by two different consultants at Pediatric Radiology, Ullevaal University Hospital, and mainly by one radiologist at Unilabs Norway. The dental age assessments have been performed by four different dentists, two for each age assessment, and another dentist has assured the quality of all the assessments.

In total 3 819 individuals have been included in the analysis, 3 333 boys and 486 girls. Figure 5 shows the number of age assessments performed from January 2010 to December 2014, separated in gender and years. There are 2 416 individuals in the first time period (January 2010 - July 2013), 2 124 boys and 292 girls. In the second time period (August 2013 - December 2014) there are 1 403 individuals, 1 209 boys and 194 girls.

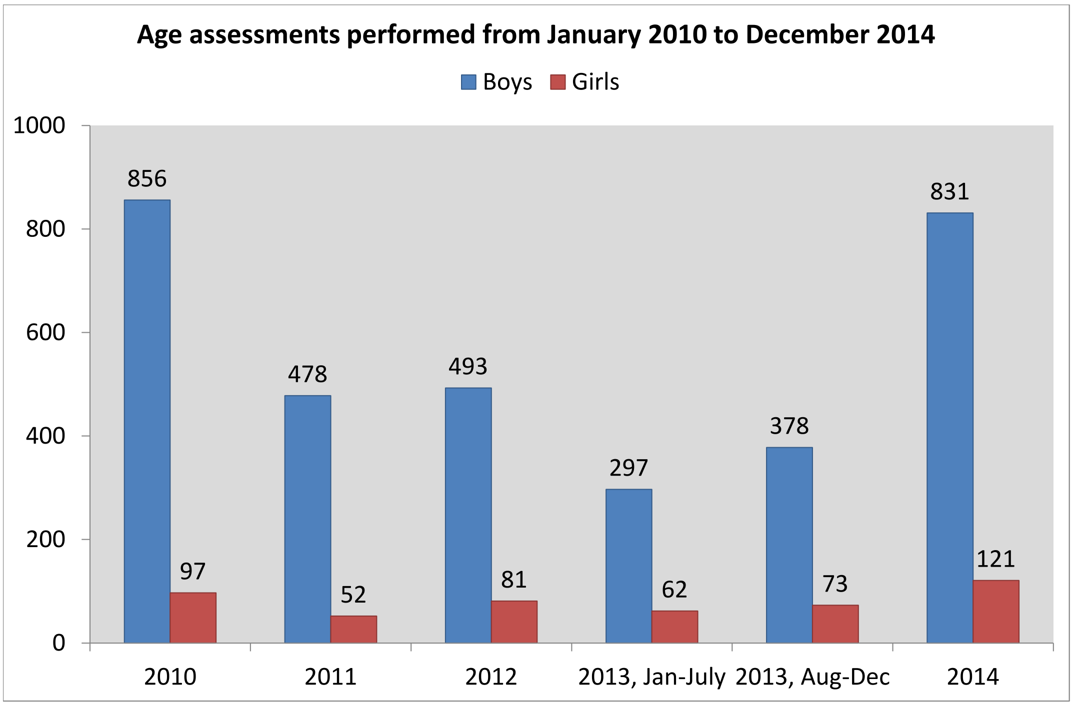


Figure 5. Number of age assessments performed from January 2010 to December 2014, separated in gender and years. Age assessments performed in 2013 consist of two periods; January to July and August to December.

A change over time in the relationship between the skeletal and dental age assessments has been observed. When the skeletal age assessment was performed by Pediatric Radiology, Ullevaal University Hospital, the skeletal age was often higher than the dental age. After Unilabs Norway has taken over the skeletal age assessment, there are more asylum seekers with a lower skeletal age than dental age.

The purpose of the analysis has been to check for real differences and reveal any systematic differences between the two independent age assessments. Variables used in the statistical analysis are: skeletal and dental age assessments (given as integers), final age assessment (given as integer), and the background variables gender and citizenship at individual level.

3.1.2 Methods

The age assessment from radiographs of the left hand and wrist is based on the atlas of Greulich and Pyle (1959), which is the most commonly used method for skeletal age estimation from hand radiographs. A radiologist manually compares the radiographs of the individual to the atlas to grade the development of bones in the hand and wrist. The skeletal age is assessed from the nearest matching reference radiograph in the atlas. The task for Unilabs Norway is to give the skeletal age assessment as an integer. According to the atlas of Greulich and Pyle, skeletal growth is complete for girls when they reach 18 years of age and for boys aged 19. When skeletal growth has ceased, a radiographic examination cannot tell anything more than that a girl is 18 years or older or that a boy is 19 years or older, and the skeletal age assessement is set to 18 years for girls and 19 years for boys.

The dental age assessment was determined with a dental clinical examination and orthopantomograms (OPGs) of the jaw, which are compared to a reference atlas of stages. Each stage of crown and root correspond to a mean or median dental age. The principle of a grading system is to read the dental age from published tables, after the root stage has been selected. In young adolescents all teeth but the third molars (wisdom teeth) have completed root development (Haavikko, 1970). The root development of mandibular third molar has been assessed from tables published by Liversidge (2008). In this paper, separate tables for males and females for four ethnic groups, two groups in UK and two groups in South Africa, were given. For individuals with missing third molars, root development was based on the second molar (Haavikko, 1970) and reduction of pulp volume of a front tooth (Kvaal et al., 1995). The frequency of missing third molars vary across populations (see Table 1 in Hentisz, 2003). 20-25% of the individuals in our analysis had missing third molars. The dental age assessments in this analysis were given as integers.

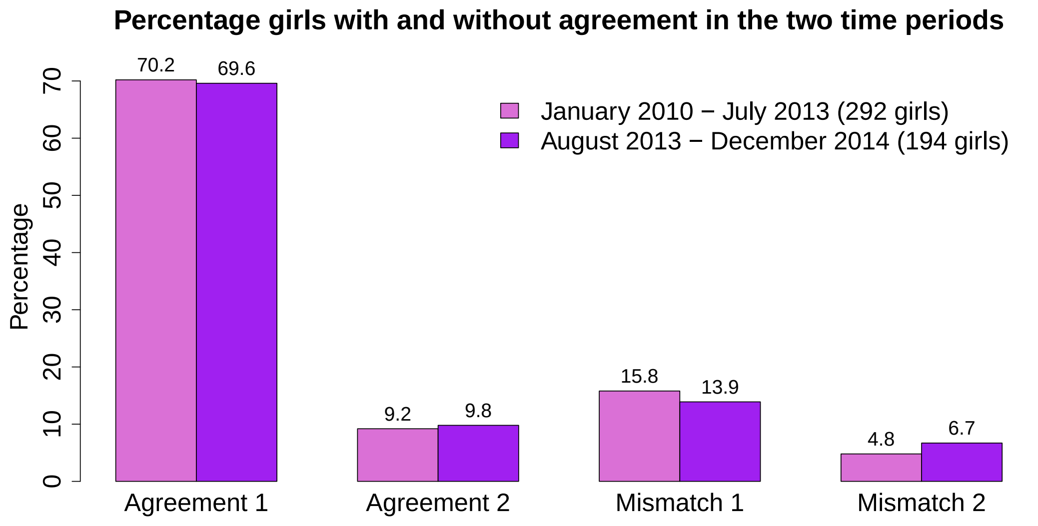
Agreement between the two age assessments is defined according to the asylum seeker's age being assessed to be above or below 18 years, with four possible outcomes:

- **Agreement 1:** The individual is 18 years or older from both of the age assessments.
- **Agreement 2:** The individual is below 18 years from both of the age assessments.
- **Mismatch 1:** The individual is 18 years or older from the skeletal age assessment and below 18 years from the dental age assessment.
- **Mismatch 2:** The individual is below 18 years from the skeletal age assessment and 18 years or older from the dental age assessment.

We have applied the Mann-Whitney U test to compare the means of two independent groups of samples that do not necessarily follow a normal distribution, it is a nonparametric test. In all hypothesis tests, the significance level is 5%. If the null hypothesis is rejected, the outcome is said to be statistically significant. The lower the significance level, the more the data must diverge from the null hypothesis to be significant.

3.1.3 Results

From Figure 6 we see that the two age assessments are in agreement for 79.5% and 79.4% of the girls for the two data periods, and for 86.3% and 77.9% of the boys. The difference between the percentage boys with agreement is -8.4 with a 95% confidence interval from -11.2 to -5.5. The significant reduction for boys reflects that the percentage of asylum seekers with skeletal age less than 18 years and dental age 18 years or older (Mismatch 2) has increased from 4.0% in the first data period to 12.4% in the second data period.



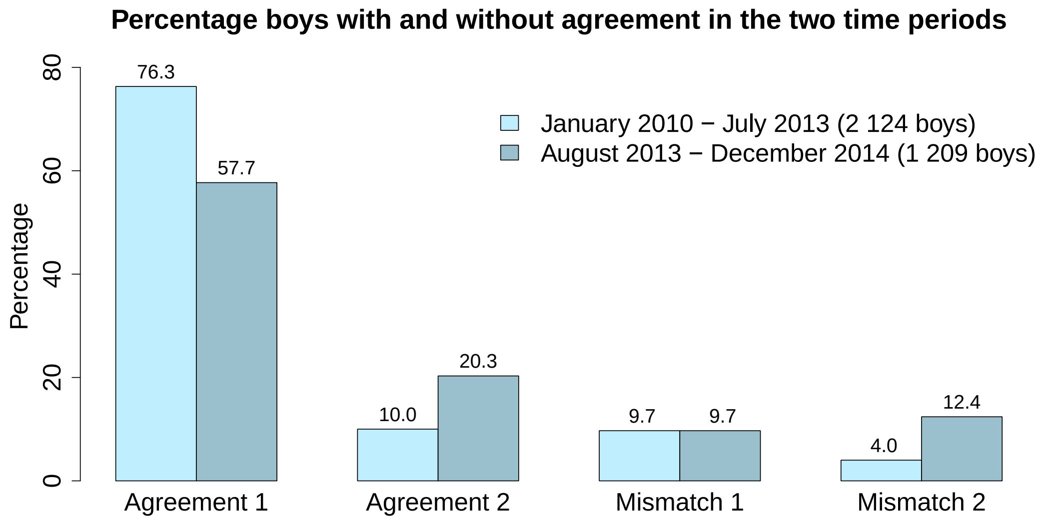


Figure 6. Percentage girls (top) and boys (bottom) with and without agreement for age assessments performed from January 2010 to July 2013 (292 girls and 2 124 boys) and from August 2013 to December 2014 (194 girls and 1 209 boys). Agreement is defined according to the asylum seeker's age being assessed to be above or below 18 years.

From a Mann-Whitney U test, the skeletal age assessment is significantly higher than the dental age assessment in both data periods for girls so there has been no change over time in the relationship between the two age assessments for girls. From a Mann-Whitney U test, the skeletal age assessment is significantly higher than the dental age assessment in the first data period for boys. In the second data period, there is a larger proportion of boys with a lower skeletal age assessment than dental age assessment. Although there is no significant difference between the two age assessments in the second data period, there has been a change over time in the relationship between the two age assessments for boys.

There are several factors that may explain the change over time in the relationship between the skeletal and dental age assessments for boys. There is a difference in the age- and citizenship distributions for boys in the two data periods. Figure 7 shows the number of final age assessments from BarneSak AS in ten different age groups in the two time periods, which we have used to look at the age distribution. The last group (≥ 20) must be interpreted as 20 years or older. 1 645/2 124 $\approx$ 77.4% of the boys in the first time period has a final age assessment of 18 years or older, compared to 716/1 209 $\approx$ 59.2% in the second time period. A Mann-Whitney U test shows that the final age assessment is significantly higher in the first data period.

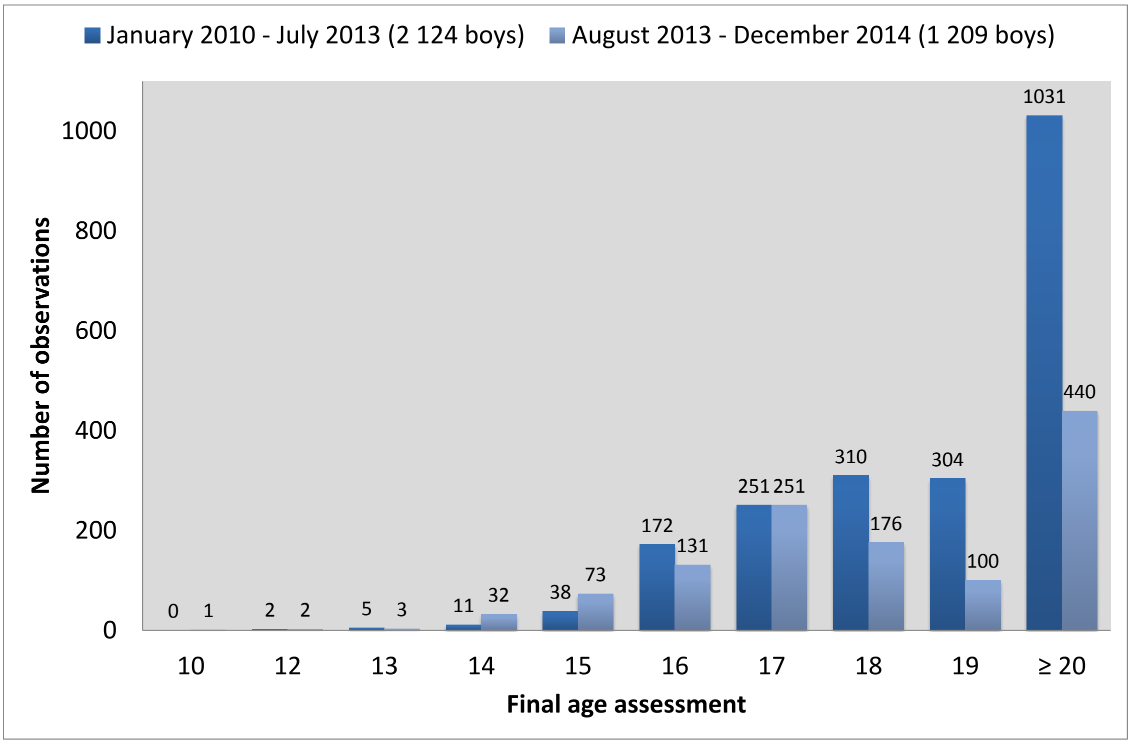


Figure 7. Final age assessments from BarneSak AS in the two time periods, separated in ten different age groups. The last group (≥ 20) must be interpreted as 20 years or older.

There are different groups of unaccompanied minor asylum seekers in the two data periods. The majority of the age assessments has been performed on Afghan, Eritrean and Somali boys in both data periods (81.7% and 80.4%, respectively). Between the two data periods, the percentage of age assessments performed on Afghan, Eritrean and Somali boys has respectively been considerably decreased (from 59.6% to 29.4%), considerably increased (from 4.5% to 31.2%) and almost unchanged (17.7% and 19.8%). Table 1 gives the percentage of boys where the two age assessments are in agreement for the two data periods, for all citizenships together and separately for Afghanistan, Eritrea and Somalia. The table shows that the reduction in agreement in the second data period is especially true for boys from Afghanistan and Eritrea. We consider the Somali boys to be in good agreement in both data periods.

Table 1. Percentage boys where the skeletal and dental age assessments are in agreement for the two data periods, for all citizenships together and separately for Afghanistan, Eritrea and Somalia. Agreement is defined according to the asylum seeker's age being assessed to be above or below 18 years.

	All	Afghanistan	Eritrea	Somalia
January 2010 - July 2013	86.3%	84.6%	80.2%	89.9%
August 2013 - December 2014	77.9%	76.1%	69.0%	89.5%

3.1.4 Conclusions

We consider it to be good agreement between the age assessments based on skeletal maturity and the age assessments based on dental development performed on girls and boys in Norway from 2010 to 2014, according to the asylum seeker's age being assessed to be above or below 18 years.

These data provide no basis to conclude that Pediatric Radiology, Ullevaal University Hospital, and Unilabs Norway interpret radiographs of the left hand and wrist significantly different. Firstly, a change over time in the relationship between the skeletal and dental age assessments only applies to boys. In this context it is important to emphasize that there are substantially fewer girls than boys in the analyzed data (respectively 12.1% and 13.8% of the total number of individuals in the two data periods). Secondly, we have found a difference in the age- and citizenship distributions for boys in the two data periods. A shift in relationship between the two age assessments for boys may be explained by a younger age group in the second data period and different groups of unaccompanied minor asylum seekers in the two data periods.

3.2 Comparison of manual and automated skeletal age assessment

3.2.1 Background

Haugen et al. (2016) consider it to be good agreement between the skeletal and dental age assessments performed in Norway from 2010 to 2014, according to the asylum seeker's age being assessed to be above or below 18 years. Nevertheless, visual grading techniques have several drawbacks. Firstly, variability related to subjectivity and secondly, inter- and intra-observer differences. The purpose of automation is to make the age assessment more objective and less dependent on the radiologists performing the visual analysis.

Unilabs Norway has tested BoneXpert, a software for automated estimation of skeletal age from a child's hand radiograph. The test has been performed on a sample of 100 left hand and wrist radiograhps of boys. A manual skeletal age assessment of these boys has been performed by Unilabs Norway in the period from August 2013 to December 2014. The purpose of this study has been to understand agreement and differences between the two skeletal age assessments. To ensure that the test is performed on relevant individuals, 95 of the boys in the sample have a final age assessment of 15 years or higher and 67 boys have a final age assessment of 17 years or higher.

3.2.2 Methods

The age assessment from radiographs of the left hand and wrist is based on the atlas of Greulich and Pyle (1959). The task for Unilabs Norway is to give the skeletal age assessment as an integer. According to the atlas of Greulich and Pyle, skeletal growth is complete for boys when they reach 19 years of age. When skeletal growth has ceased, a radiographic examination cannot tell anything more than that a boy is 19 years or older and the skeletal age assessment is set to 19 years.

BoneXpert is a software for automated estimation of skeletal age from a child's hand radiograph based on methods for image analysis of radiographs of the hand and carpus, developed in Denmark by the CEO of Visiana, Hans Henrik Thodberg, and others. The software is described in Thodberg (2009) and Thodberg et al. (2009), and is well tested in several large studies for Greulich and Pyle skeletal age 2.5-17 years for boys and 2-15 years for girls (Martin et al., 2013; Thodberg and Sävendahl, 2010; van Rijn et al., 2009). BoneXpert has a continuous scale so the skeletal age is given with two decimal digits.

Unilabs Norway has tested a prototype of BoneXpert developed on the basis of 231 radiographs of healthy Swiss children born in 1954-1956. Greulich and Pyle skeletal age is 2.5-19 years for boys. For boys with skeletal age greater than 17 years, only the radius and ulna bones are still under development and can be used for skeletal assessment. Reliable localization and interpretation of the radius bone is crucial for the software for boys who are older than 18 years, since the ulna bone then is fully developed.

Agreement between the two skeletal age assessments is defined according to the asylum seeker's age being assessed to be above or below 18 years, with four possible outcomes:

- **Agreement 1:** The individual is 18 years or older from both of the skeletal age assessments.
- **Agreement 2:** The individual is below 18 years from both of the skeletal age assessments.
- **Mismatch 1:** The individual is 18 years or older from the manual skeletal age assessment and below 18 years from the automated skeletal age assessment.
- **Mismatch 2:** The individual is below 18 years from the manual skeletal age assessment and 18 years or older from the automated skeletal age assessment.

There is no operational standard to determine whether the individual is above or below 18 years. The manual skeletal age assessment is given as an integer and the individual is therefore 18 years or older if skeletal age is 18 years or higher. The continuous scale of BoneXpert provides more flexibility in the choice of a criterion. Due to biological variations, we will never get error rates for Mismatch 1 and Mismatch 2 equal to zero. However, a decision rule can be tuned to balance the error rates to acceptable levels. Two different criteria have been analyzed, where the second criterion probably is better suited to a continuous scale:

- **Criterion 1**: The skeletal age from BoneXpert must be 18 years or higher for the individual to be above 18 years.
- **Criterion 2**: The skeletal age from BoneXpert must be 17.5 years or higher for the individual to be above 18 years.

3.2.3 Results

Figure 8 shows a scatter plot of the automated skeletal age assessment versus the manual skeletal age assessment for a test sample of 100 radiographs of boys, together with the line of equality (given in red). The manual skeletal age assessment is given as an integer and is separated in seven age groups from 13 years up to 19 years. Automated skeletal age assessment is given with two decimal digits. The figure shows that when the manual skeletal age assessment is in the age groups from 13 years to 16 years, the automated skeletal age assessment tend to be higher. For the age group 19 years, automated skeletal age assessment tend to be lower than the manual skeletal age assessment.

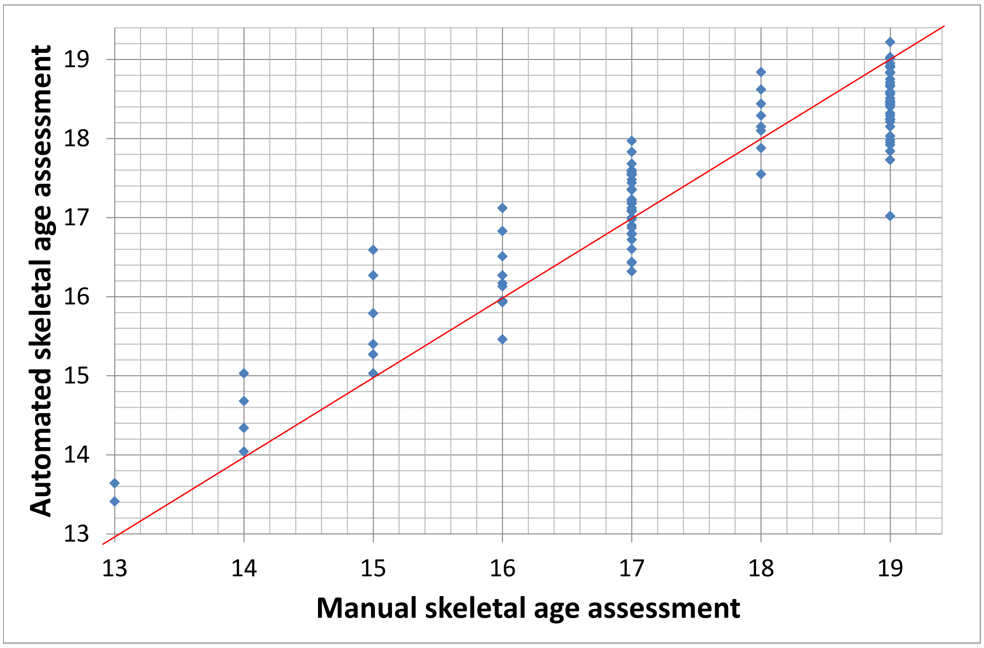


Figure 8. Scatter plot of automated skeletal age assessment versus manual skeletal age assessment for our test sample of 100 radiographs of boys. The line of equality is given in red. Manual skeletal age assessment is given as an integer. Automated skeletal age assessment is given with two decimal digits.

The root-mean-square error (RMSE) is a measure of the difference between values and the RMSE between the two skeletal age assessments in Figure 8 is 0.61 years.

Figure 9 shows that the agreement between the manual and automated skeletal age assessment performed at Unilabs Norway is 92% for criterion 1 and 91% for criterion 2 on our test sample of 100 radiographs of boys. For criterion 1, 39% of the boys are 18 years or older from both of the skeletal age assessments (Agreement 1) and 53% of the boys are below 18 years from both of the skeletal age assessments (Agreement 2). For criterion 2, the percentage of boys in the two groups with agreement is 46% and 45%, respectively.

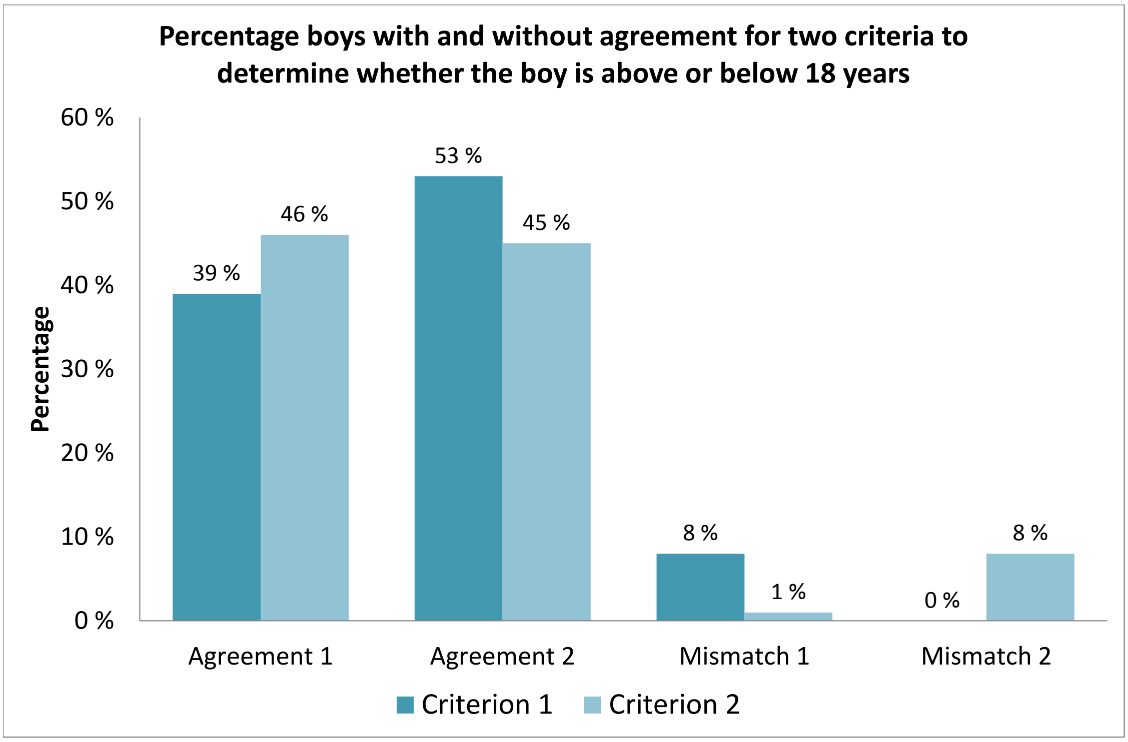


Figure 9. Percentage boys with and without agreement for a test sample of 100 radiographs of boys. Agreement is defined according to the asylum seeker's age being assessed to be above or below 18 years. The manual skeletal age assessment must be 18 years or higher for the boy to be above 18 years. Skeletal age from BoneXpert must be 18 years or higher (criterion 1) or 17.5 years or higher (criterion 2) for the boy to be above 18 years.

3.2.4 Conclusions

We have performed a successful test of BoneXpert on a sample of 100 left hand and wrist radiograhps of boys. The agreement between the manual and automated skeletal age assessment is 92% for criterion 1 and 91% for criterion 2, according to the asylum seeker's age being assessed to be above or below 18 years. We consider this to be good agreement, which is a result of a precisely performed manual age assessment, a precise and reliable automated age assessment and good quality of the analyzed radiographs. However, there is a need to verify this finding on a larger dataset.

BoneXpert needs also to be applied on a larger sample of radiographs in order to decide an operational standard to determine whether the individual is above or below 18 years based on the automated skeletal age assessment. If the criterion in our test had been that skeletal age from BoneXpert must be higher than 17.7 years, the agreement between the two skeletal age assessments had been 96%. In a larger study a percentage error related to the operational standard can be calculated and tuned to balance the error rates for Mismatch 1 and Mismatch 2 to acceptable levels.

3.3 Recommendations

For Norwegian authorities, and in particular the UDI, it is important to determine with reasonable certainty whether the asylum seeker is older or younger than 18 years of age. In this work package, agreement between two age assessments has therefore been defined according to the asylum seeker's age being assessed to be above or below 18 years.

The development of hand bones and teeth are supposed to be biologically independent (Gelbrich et al., 2015) and are recommended used in age assessment procedures (Schmeling et al., 2008). We consider it to be good agreement between the age assessments based on skeletal maturity and the age assessments based on dental development performed on girls and boys in Norway from 2010 to 2014 (section 3.1). When there is a mismatch in the two age assessments, we recommend the UDI to use more information about the asylum seeker in the age assessment. These are the cases where:

- The individual is 18 years or older from the skeletal age assessment and below 18 years from the dental age assessment.
- The individual is below 18 years from the skeletal age assessment and 18 years or older from the dental age assessment.

Automated methods with continuous age will allow for more precise description of variation and uncertainty. These methods can make the age assessment more objective and less dependent on the professionals performing the visual analysis. The results from the test of BoneXpert in section 3.2 are very promising. Prior to implementing BoneXpert in the age assessment process in Norway, this finding needs to be verified on a larger sample of radiographs. The automated skeletal age assessment is given with two decimal digits and a larger study is needed to decide an operational standard to determine whether the individual is above or below 18 years based on this assessment. Due to biological variations, we will never get error rates for the two outcomes of mismatches between the two skeletal age assessments equal to zero. However, a decision rule can be tuned to balance the error rates to acceptable levels.

3.4 References

Gelbrich B, Frerking C, Weiss S, Schwerdt S, et al. (2015). Combining wrist age and third molars in forensic age estimation: how to calculate the joint age estimate and its error rate in age diagnostics. Ann Hum Biol., Vol. 42(4):389-396.

Greulich WW, Pyle SI. (1959). Radiographic Atlas of Skeletal Development of the Hand and Wrist. Stanford, CA: Stanford University Press, 2nd edition.

Haavikko K. (1970). The formation and the alveolar and clinical eruption of the permanent teeth. An orthopantomographic study. PhD thesis, University of Helsinki, Finland.

Haugen M, Grøgaard J, Kjellevand TO, Kvaal SI. (2016). Sammenligning av to metoder for aldersvurdering av enslige mindreårige asylsøkere. NR-Note, SAMBA/05/16.

Haugen M, Grøgaard J, Kjellevand TO. (2016). Sammenligning av manuell og automatisert skjelettvurdering. NR-Note, SAMBA/13/16.

Kvaal SI, Kolltveit KM, Thomsen IO, Solheim T. (1995). Age estimation of adults from dental radiographs. Forensic Sci Int., Vol. 74(3):175-185.

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

Martin DD, Heil K, Heckmann C, Zierl A, et al. (2013). Validation of automatic bone age determination in children with congenital adrenal hyperplasia. Pediatr Radiol., Vol. 43(12):1615-1621.

Schmeling A, Grundmann C, Fuhrmann A, Kaatsch HJ, et al. (2008). Criteria for age estimation in living individuals. Int J Legal Med., Vol. 122(6):457-460.

Thodberg HH. (2009). Clinical review: An automated method for determination of bone age. J. Clin. Endocrinol. Metab, Vol. 94(7):2239-2244.

Thodberg HH, Kreiborg S, Juul A, Pedersen KD. (2009). The BoneXpert method for automated determination of skeletal maturity. IEEE Trans Medical Imaging, Vol. 28(1):52-66.

Thodberg HH, Sävendahl L. (2010). Validation and reference values of automated bone age determination for four ethnicities. Acad Radiol., Vol. 17(11):1425-1432.

van Rijn RR, Lequin MH, Thodberg HH. (2009). Automatic determination of Greulich and Pyle bone age in healthy Dutch children. Pediatr Radiol., Vol. 39(6):591-597.

4 WP2: Better understanding of dental variations between groups

The objective of this activity has been to identify and quantify differences in tooth development between ethnic groups and genders and how this impacts on estimated age from teeth[3]. To obtain this, age distributions for molar stages have been derived and compared across ethnic groups and genders. We have provided an approach for deriving age distributions conditional on molar stage by separate analyses for males and females for five ethnic groups. The analysis has been performed separately on scored data for mandibular second molar and mandibular third molar (wisom tooth).

> The following summarizes the work that has been done within WP2, focusing on the analyses for males and females on the mandibular third molar stages 10 (root one quarter) through 14 (apex half closed). Results for other stages and for the mandibular second molar can be found in the more technical in-depth note: «Age distribution estimation given molar stages for males and females for five ethnic groups» (Tvete et al., 2016).

Helen Liversidge has provided the data, which consist of individuals with known chronological ages and molar stage scores, were the scores were made from orthopantomograms (OPGs) of the lower jaw (mandible). The scoring was performed in collaboration with international colleagues. Lyle W. Konigsberg, Professor at University of Illinois at Urbana-Champaign, has provided the statistical method and helpfully shared R and WinBUGS code. Ingunn Fride Tvete has conducted the statistical analysis and written the note.

4.1 Background

The development of a tooth can be divided into a series of discrete maturity stages by comparing OPGs of the jaw to a reference table of stages. Each stage corresponds to a mean or median dental age, depending on the reference table. Once the stage has been selected, dental age can be read from published tables, which is the principle of a grading system. Mandibular third molar formation stages and descriptive criteria are given in Figure 1 in Liversidge (2008).

This applies to all 32 teeth and the age of a child is derived from the average of the mean or median tooth ages. In young adolescents all teeth but the third molars have completed root development. The root development of mandibular third molar has been assessed from tables published by Liversidge (2008). In this paper, separate tables for males and females for four ethnic groups, two groups in UK and two groups in South Africa, were given.

3. We emphasize that the statistical analysis in WP2 separates from the dental age assessment performed on unaccompanied minor asylum seekers in Norway, as described in section 3.1.2. We have not considered how the dental age assessment of individuals is achieved in practice in this activity. WP2 provides a mathematical framework for estimating age distributions given molar stages for males and females for five ethnic groups.

The data for this analysis consisted of 1 974 males and 2 456 females with known chronological age, in the age range 10-26 years, from five ethnic groups and scores from third molar stages (left mandibular teeth). Table 2 displays the number of males and females in each of the five ethnic groups.

Table 2. Number of observations for males and females for the five different ethnicities for the third molar.

	Sub-Saharan African	Japanese	Malaysian	White/ European UK	Bangladeshi UK
Males	649	193	402	413	317
Females	718	307	554	579	298

Figures 10 and 11 show the developmental stage for the third molar versus chronological age for males and females, respectively, for the five ethnic groups. There was a great deal of overlap between ages across the molar stages. The upper right plot in Figure 10 shows that there were few observations of Japanese males, especially in the later molar stages, and this ethnic group was therefore excluded from the analysis (this was just the case for the third molar data for males, not the second molar data for males or the second molar and third molar data for females).

4.2 Methods

The estimated age given molar stage distributions were obtained in a two-step procedure, as given by Lyle W. Konigsberg. In the first step a probit regression model for each ethnic group was fitted to the mandibular third molar data. In the model, the molar stage was assumed dependent upon the natural logarithm of the known age. The purpose of the probit model was to estimate the mean age of transition between the molar stages and the common standard deviation for transitions (on log scale). In the second step these estimates were given (assumed known) in a Bayesian model for deriving the estimated age distributions for each molar stage. The ethnic groups were considered separately.

In the latter step the age for a given molar stage was a priori uncertain, and this uncertainty was expressed by letting age be uniformly distributed between 0 and 110 years. Fitting this model gave the estimated age distribution given molar stage (the so-called posterior age distribution given molar stage). The estimated age distribution for the earliest and latest molar stage will be bounded by the limits in the a priori uniform distribution (i.e. the earliest stage by 0 and the latest stage by 110), and is therefore not reported (stage five is anyway the earliest stage considered for the mandibular third molar data).

Step one was carried out using the statistical software R (R Development Core Team, 2015) and step two was carried out using WinBUGS, a statistical software for Bayesian analysis using Markov chain Monte Carlo (MCMC) methods (Lunn et al., 2000). In step two 2 500 ages were sampled (from the posterior age distribution given molar stage), forming the estimated age distribution.

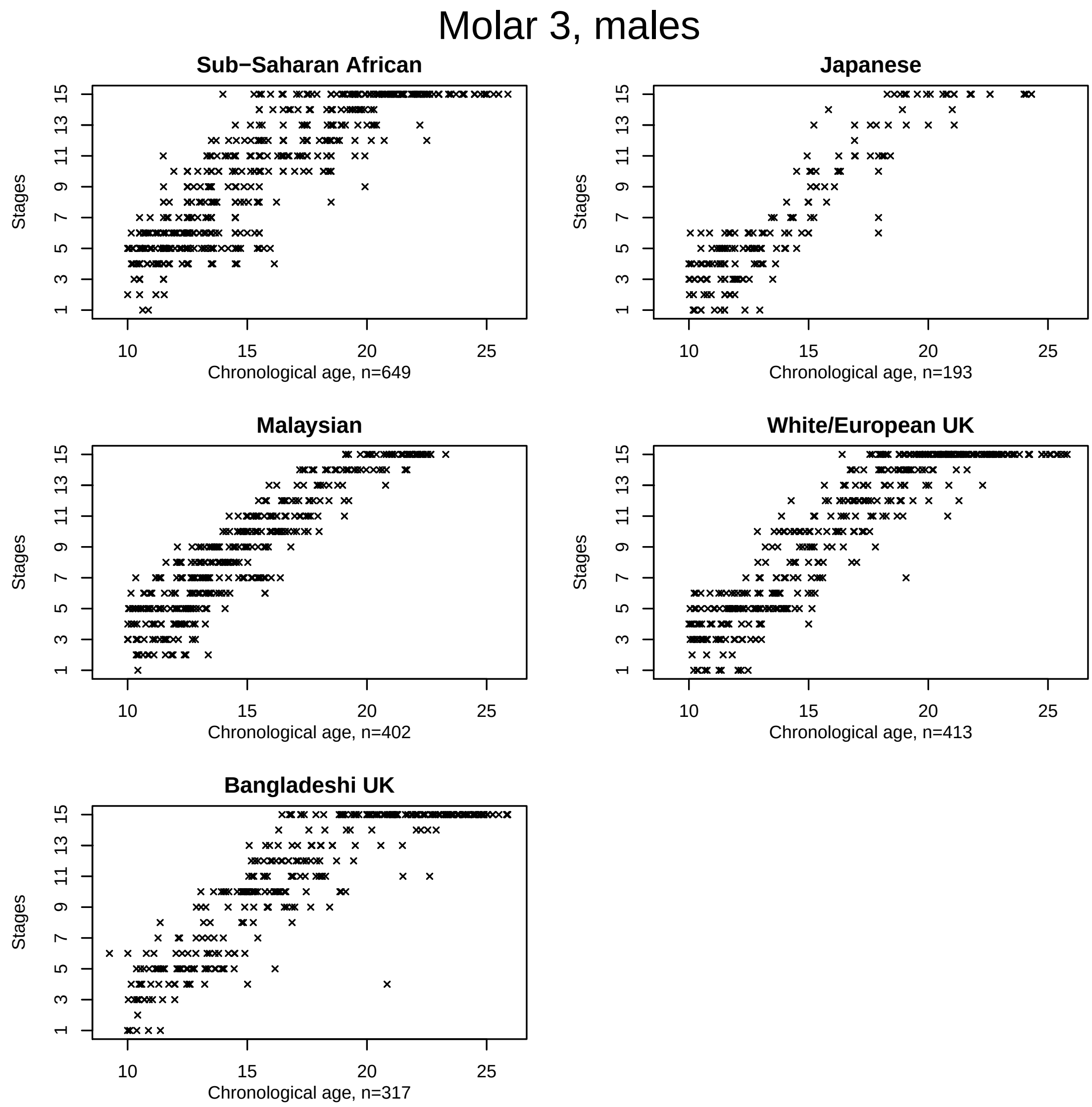


Figure 10. The developmental stage for the third molar versus chronological age for males. The data points have been jittered to avoid overplotting. Clusters of data points indicate several observations.

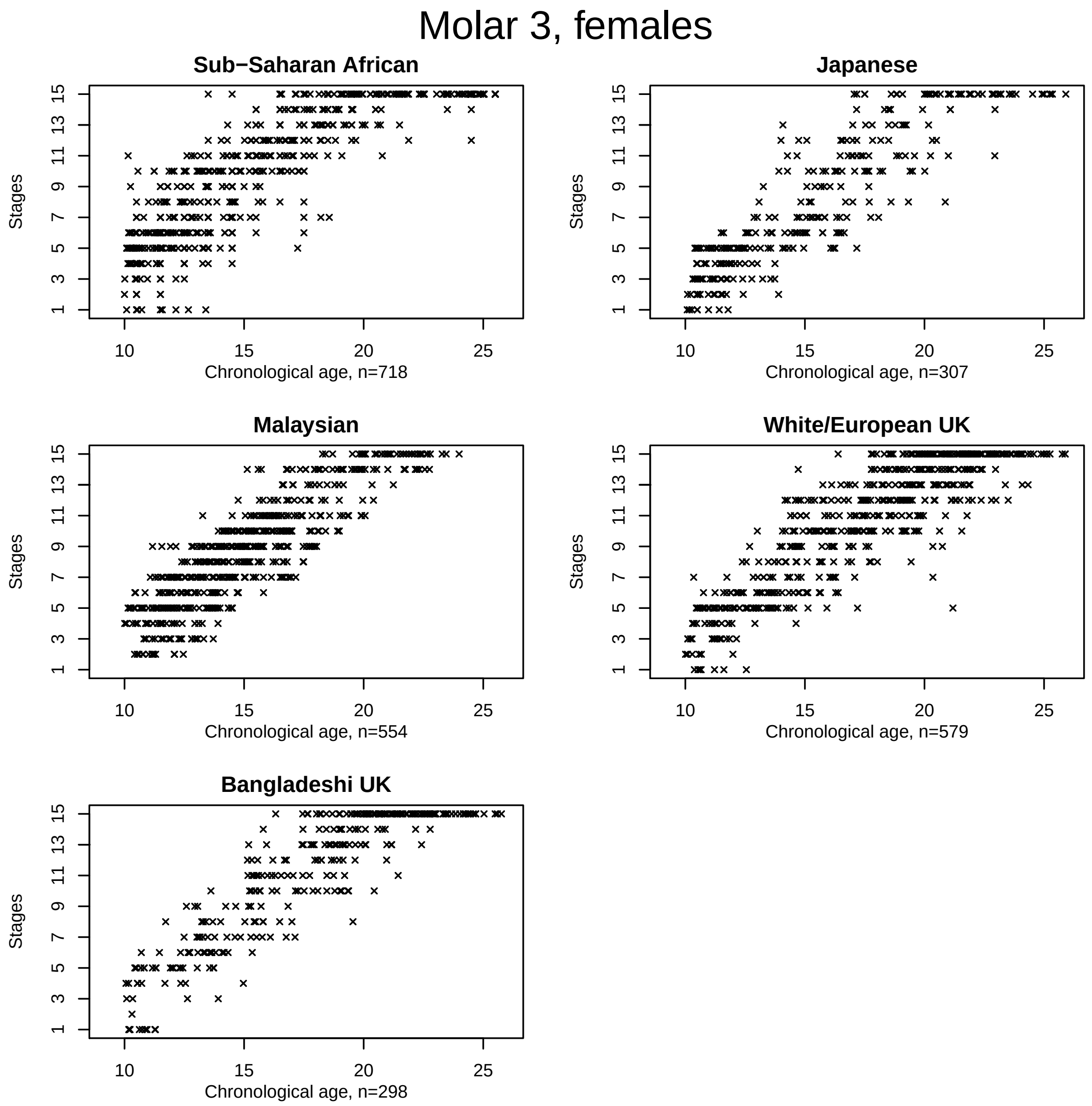


Figure 11. The developmental stage for the third molar versus chronological age for females. The data points have been jittered to avoid overplotting. Clusters of data points indicate several observations.

4.3 Results

Figures 12 and 13 display the estimated age distributions, given stage 10-14 for mandibular third molar for males and females, respectively. Tables 3 and 4 give the estimated mean age and associated 95% uncertainty intervals in years for mandibular third molar stages 10-14 for males and females, respectively, together with the estimated percentage of males/ females being 18 years or older (all derived from the 2 500 age samples obtained in step two in the estimation procedure).

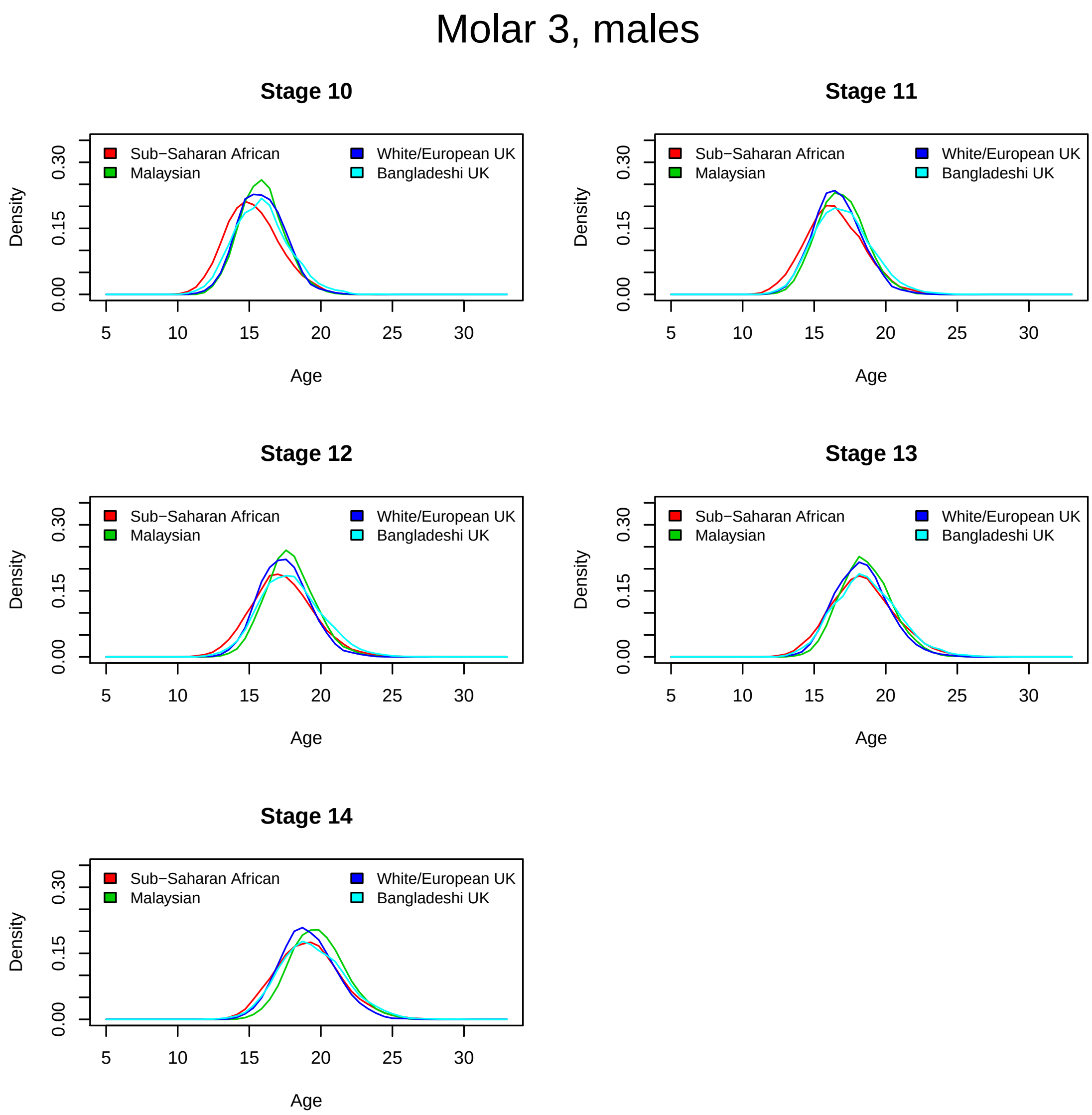


Figure 12. Estimated age distributions, given stages 10-14 for mandibular third molar for males.

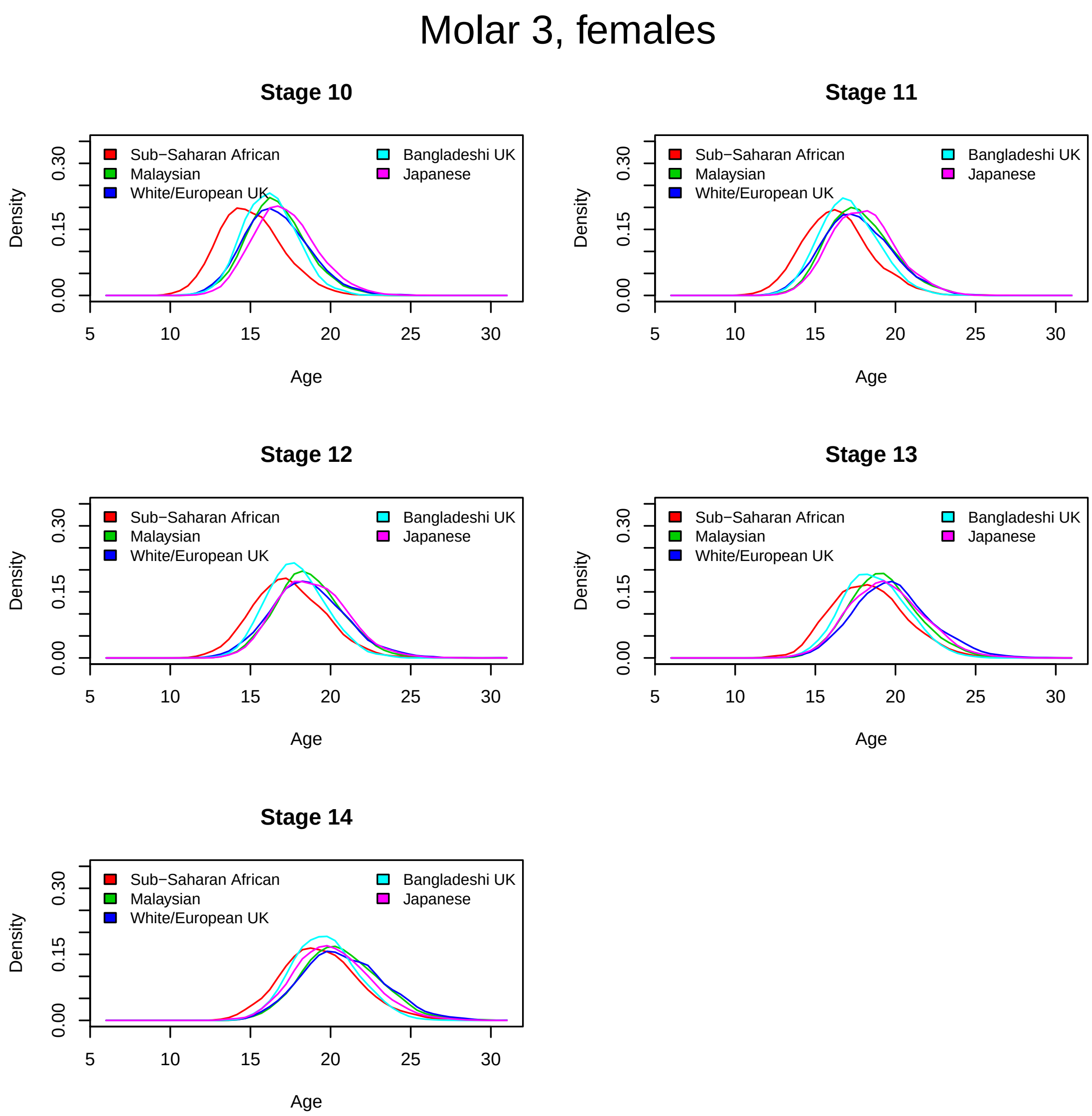


Figure 13. Estimated age distributions, given stages 10-14 for mandibular third molar for females.

Table 3. Estimated mean age and 95% uncertainty intervals in years for mandibular third molar stages for males, together with the estimated percentage of males being 18 years or older.

Stage	Sub-Saharan African	Malaysian	White/ European UK	Bangladeshi UK
10	15.26 (11.95,19.25) 8.44%	15.88 (13.10,19.10) 8.68%	15.86 (12.93,19.04) 8.88%	15.86 (12.53,19.94) 13.44%
11	16.46 (12.81,20.61) 21.68%	16.89 (13.91,20.34) 24.56%	16.65 (13.67,19.93) 20.48%	17.00 (13.57,21.25) 28.84%
12	17.36 (13.39,21.89) 36.00%	17.85 (14.82,21.52) 44.16%	17.45 (14.29,20.95) 36.32%	17.90 (14.04,22.43) 45.68%
13	18.43 (14.37,23.18) 55.72%	18.62 (15.46,22.27) 62.08%	18.32 (15.04,22.29) 56.16%	18.67 (14.67,23.46) 60.40%
14	19.16 (15.17,23.98) 68.04%	19.76 (16.36,23.65) 82.96%	19.10 (15.62,23.08) 71.16%	19.42 (15.45,24.20) 72.80%

Table 4. Estimated mean age and 95% uncertainty intervals in years for mandibular third molar stages for females, together with the estimated percentage of females being 18 years or older.

Stage	Sub-Saharan African	Malaysian	White/ European UK	Bangladeshi UK	Japanese
10	15.04 (11.62,19.25) 7.84%	16.66 (13.19,20.64) 22.08%	16.66 (12.99,21.00) 24.00%	16.24 (13.13,19.71) 14.60%	17.16 (13.72,21.43) 31.56%
11	16.40 (12.66,20.88) 20.60%	17.79 (14.39,22.12) 42.36%	17.63 (13.77,22.17) 40.80%	17.14 (13.96,20.91) 30.44%	18.01 (14.43,22.20) 49.68%
12	17.43 (13.41,22.09) 37.44%	18.7 (14.95,23.10) 61.40%	18.63 (14.35,23.66) 59.00%	17.89 (14.68,21.63) 45.44%	18.85 (15.09,23.39) 62.68%
13	18.38 (15.02,24.67) 54.52%	19.29 (15.44,23.75) 72.12%	19.82 (15.57,24.95) 77.56%	18.75 (15.17,22.80) 62.12%	19.46 (15.34,24.18) 72.4%
14	19.42 (15.02,24.67) 71.36%	20.68 (16.41,25.73) 88.12%	20.79 (16.24,26.12) 87.12%	19.68 (15.99,23.77) 80.04%	20.15 (16.01,25.04) 82.60%

A probit model is a type of regression model where the dependent variable can only take certain discrete values, such as molar stages. Under the first step described in section 4.2, different mean ages of transition between the molar stages and different standard deviations for transitions for the ethnic groups were assumed in the probit model. We call this model 1.

To examine whether there were differences between ethnic groups, we have, in addition to seperate models for all groups (model 1), considered a joint model (model 2) and a semi-joint model (model 3) for all ethnic groups in the first step of the two-step estimation procedure. Model 2 had a common slope and a common intercept for the ethnic groups. In model 3 one parameter was specified to be different for all groups, see Tvete et al. (2016) for details. To evaluate these models, the following three model evaluation approaches were considered:

1. The Bayesian information criterion (BIC),
2. Akaikes information criterion (AIC) and
3. Log-likelihood ratio test.

Model 3 performed best according to the BIC criteria. According to the AIC criteria and the log-likelihood ratio test model 1 (separate models for each ethnic group - our fitted model) performed best. All three model evaluation criteria indicated a difference between ethnic groups. This was true for both males and females.

Testing for differences between genders for each ethnic group revealed less uniform results over the model evaluation criteria, and the degree of difference was overall smaller across gender than across ethnic groups. Still, overall there were indications of gender differences.

It must be kept in mind that these tests revealed differences between ethnic groups, but do not say anything about where the differences lie. More detailed ("head-to-head") comparisons between ethnic groups need to be done in order to determine which ethnic groups that differ.

4.4 Conclusions

This work dealt with deriving age distributions for males and females given mandibular third molar stages and comparing these across ethnic groups and genders. We have provided a mathematical framework for doing this. When comparing different probit regression models, three model evaluation criteria concluded that there were differences across the ethnic groups and also to some degree between gender.

The age span in the estimated age distributions for mandibular third molar stages for males and females reflects the biological variation which will always be present when the maturation stage of physical properties is explored.

4.5 Recommendations for further research

In the second step in the age distribution procedure the parameter estimates from the probit model are considered known, and uncertainty in these parameter estimates is not taken into consideration. If these uncertainties had been taken into consideration, the estimated age distributions could be somewhat different.

The mandibular second molar and mandibular third molar data were analyzed separately. A joint modelling approach is a possibility but will give a far more complex model.

4.6 References

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

Lunn DJ, Thomas A, Best N, Spiegelhalter D. (2000). WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. Stat Comput., Vol. 10(4):325-337.

R Development Core Team (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Tvete IF, Konigsberg LW, Liversidge HM. (2016). Age distribution estimation given molar stages for males and females for five ethnic groups. NR-Note, SAMBA/12/16.

5 WP3: Analysis of MRI-based data

The objective of this activity has been to investigate further the MRI-based methods suggested by Professor Ernesto Tomei and his colleagues at the University of Rome, and to develop and explore methods based on automatic image analysis that can help in the process of analyzing these images. To obtain this, the following subtasks were defined:

(i) Obtain dataset from individuals of known age that can be exchanged between the partners of the project.

(ii) Formalize manual procedure: The manual procedure currently used should be formalized and described in a way that makes it possible to develop a more automated image-based approach for steps in the procedure.

(iii) Select steps for automation: From the manual procedure, the steps which are suitable for automated image analysis will need to be identified.

(iv) Develop methods for image analysis: Methods for the identified steps should be developed and implemented in a simple solution suitable for small scale testing.

(v) Evaluation: Evaluate and test the methods.

The following summarizes the work that has been done within WP3. More technical details, on both the methods that have been studied and those that have been developed, can be found in the more technical in-depth note (written in Norwegian): «Bruk av MR-bilder for estimering av skjelettalder» (Eikvil, 2016).

The work in this work package has been carried out at the Norwegian Computing Center (by Line Eikvil) and at the University of Rome. NR has worked on deriving an automatic image-based approach for bone age estimation from MR images, and implemented and tested this approach on MRI datasets. The University of Rome has developed the manual approach which has been the starting point for the development of the automatic method. Through the project the University of Rome has prepared and made available MRI datasets, helped in the understanding of their methods through meetings in Oslo and e-mail correspondence, prepared additional information needed for testing and validation, in addition to further adaptation and testing of the manual approach. Milvia Martino from the University of Rome has contributed especially to this with expert knowledge and work on the manual method.

5.1 Background

Bone age estimation from the hand is today based on X-ray images. Typically the bone age estimate is determined by visually comparing the X-ray image to a standard atlas, usually the atlas of Greulich and Pyle (1959). Alternatively, the development stage of each bone is determined individually by a similar comparison and the age estimate is derived from this, as suggested by Tanner and Whitehouse (1975). It is however desirable to derive approaches using nonionizing radiation, and this could also make it easier to collect new and larger reference datasets as stronger restrictions apply to collection of X-ray based data for research purposes (Serinelli et al, 2014). MR imaging could be such a nonionizing approach.

Currently there are no established methods for doing bone age estimation from MR images, but the University of Rome has been working on developing such an approach similar to the approaches used for X-ray. For manual approaches, there is however always some variability related to subjectivity and there will be inter- and intra-observer differences. Analysis of the MR-images can be particularly challenging as a lot more details are visible here than in the X-ray images. Hence, it is desirable to be able to automate parts of the analysis.

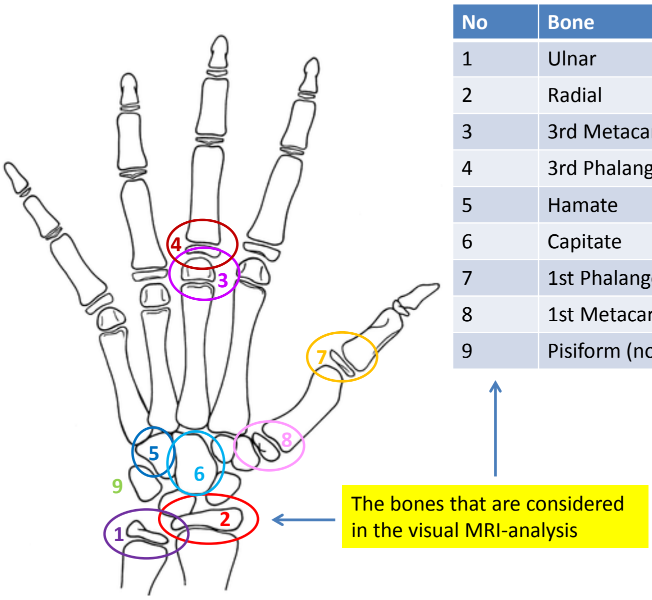
The objective of this study has been to look into whether and how a method for automatic bone age estimation can be developed based on the manual MRI-approach that has been developed at the University of Rome. To do this we have studied both the details of the manual MRI-approach as well as automatic approaches that have been suggested for X-ray and some very recent studies on automatic approaches for MRI. Based on this we have suggested a possible automatic approach, developed methods and implemented and tested these for bone age estimation based on the radial bone of the hand.

Tests have been performed on an MRI dataset provided by the University of Rome. The results indicate that the suggested method can be a viable approach for bone age estimation from MR images. However, the available dataset has been limited, and there is a need to validate both the manual and the automatic approach on a larger and independent dataset. In the following we describe the work that has been done in more detail.

5.2 Summary of the approach from the University of Rome

The method for MRI-based bone age estimation that has been developed at the University of Rome is, similar to X-ray methods, based on an atlas of images and descriptions of the development for the bones in the hand (Tomei et al., 2014b). For the MR images up to 12 development stages have been defined for 9 different bones in the hand (see Figure 14). For each development stage of each bone, an image is accompanied by the description of the appearance in the image.

While X-ray images of the hand will mainly show the structure of the bones, MR images will show much more details and also convey structures of different types of tissue. Hence, the descriptions of the development stages in the atlas from Rome contain much more details. The characteristics that are defined for each stage are typically connected to presence and size of cartilage, shape of the bone, distance between bones and degree of fusion based on observed lines and structures between bones. Furthermore, the MRI dataset of a hand consists not only of one image like the X-rays, but of a set of image slices through the hand (26 slices for the datasets in this study). In the manual process, the researchers in Rome select the one (or a few) slice(s) where the characteristics for the bone to be analyzed are most visible. Figure 15 gives an example from Tomei et al. (2014b) of the development of the radial bone as observed in MR-slices for seven stages.



No	Bone
1	Ulnar
2	Radial
3	3rd Metacarpal epiphysis
4	3rd Phalangeal epiphysis
5	Hamate
6	Capitate
7	1st Phalangeal epiphysis
8	1st Metacarpal epiphysis
9	Pisiform (not visible here)

Figure 14. Illustration of the bones that are used in the visual MRI approach for bone age estimation at the University of Rome.

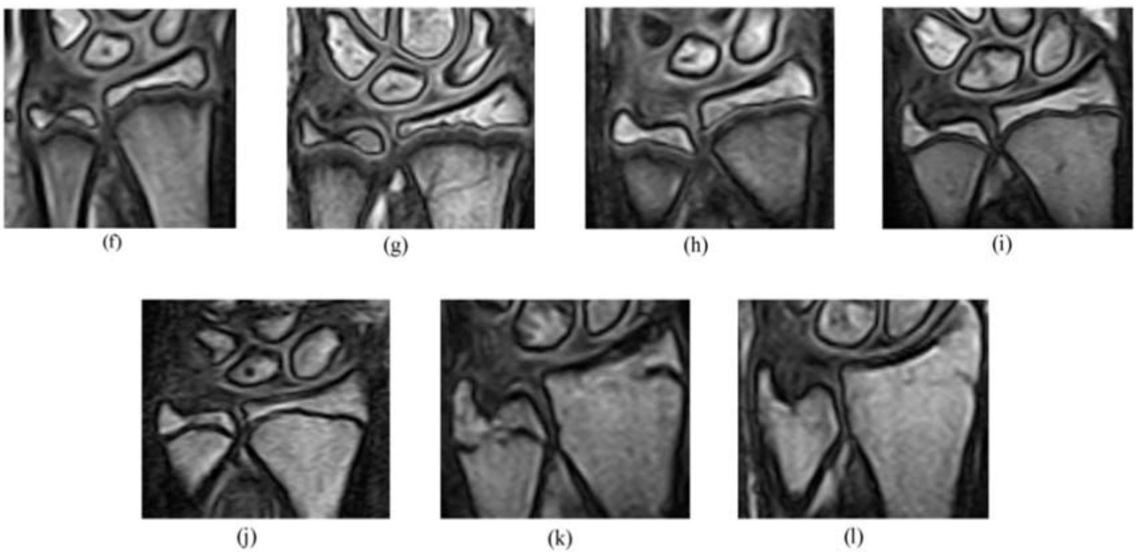


Figure 15. The development of the radial bone as observed in MRI-slices for the stages VI (f) to XII (l). Images from the University of Rome (Tomei et al., 2014b).

For a new individual the development stage for each bone is determined by comparing the selected MR-slice(s) with the images and descriptions from the atlas. This results in 9 developments stages, one for each bone, which may be different. A score is then computed as the sum of the development stages found for these 9 bones, and a look-up-table (derived through regression) defines the correspondence between this score and bone age. Figure 16 shows the correspondence between age and development stage for males as defined in the MRI-based approach from the University of Rome.

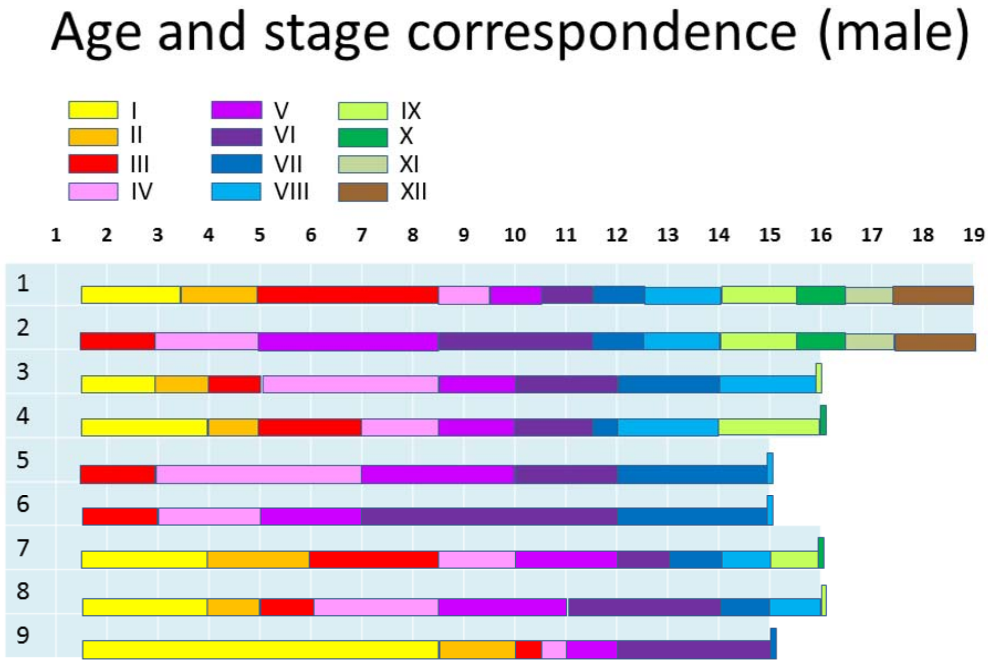


Figure 16. Correspondence between age and development stage for males as defined in the MRI-based approach from the University of Rome. Here we illustrate the development stage (colour coded, I-XII) that each bone (1-9) is expected to be in at different ages (1-19).

The dataset that has been used to establish the method consists of MRI datasets acquired from roughly 100 males and 100 females (126/106) distributed over the ages 4-18 years (Tomei et al., 2014a). Hence, the dataset is quite limited. The score model and the regression have been derived from this dataset. The method works well on this set, but has to our knowledge not yet been validated on an independent dataset or by independent radiologists (i.e. radiologists not involved in the research).

5.3 Summary of existing studies on automated bone age estimation

All automatic approaches for bone age estimation up until 2014 analyze X-ray images. In 2014 and 2015 a few studies from a group in Graz in Austria also address the issue of automatic bone age estimation from MR images (Ebner et al., 2014; Stern et al., 2014a,b; Urschler et al., 2015). The main principles are similar for X-ray and MRI. Two main tasks need to be solved to do automatic bone age estimation from images: (i) The region (or the bones) to be analyzed needs to be identified and localized automatically in the image, and (ii) the characteristics of the image region that are related to age need to be determined and extracted. The different approaches solve these tasks in different ways.

For the first task, there are in general two types of regions in the image that are localized and used for further analysis: either the contours of the bones or a region of interest (ROI) surrounding the epiphysis and/or metaphysis. Some approaches combine this as the contours of the bones are found first to help localize the regions of interest afterwards.

For the second task, most automated approaches do not use features that are based directly on characteristics that would have been chosen if the purpose was to do a visual manual analysis. Still the features can be somewhat related to these as they describe characteristics like edges, texture or distances in the image. The more recent studies do however tend to use either the image information of a region directly or very general features, rather than features specifically designed for the task. The few studies that analyze MR images go even further in this direction as they let the system select the features automatically. This is also the general trend in image analysis when it comes to recognition in complex images, where techniques based on so-called Deep Learning have shown a performance in terms of recognition rates that has not been seen before (Krizhevsky et al., 2012).

Many of the suggested automatic methods are based on the same principle as the manual approach defined by Tanner and Whitehouse (1975), where they first do classification into predefined development stages. Afterwards they derive a bone age estimate by using a model (often regression based) for the correspondence between development stage and age. The advantage of this approach is that the regression models are easily exchangeable and different correspondence models can be created for different groups.

Another approach is to estimate the bone age directly without going through development stages. The advantage of this approach is that there is no information loss through the intermediate step of quantization (classification to stages). The disadvantage is that bone age is then linked directly to the image features, and while development stages may be the same for different groups (males, females, different ethnicities), the link to bone age is not. Hence, this will require image-based models trained for each group, which may again require more image data.

5.4 Assessment of the steps with respect to automation

We have done an assessment of the various steps in the manual MRI-based process with respect to what is feasible when using an automatic approach. Based on this we have chosen a strategy for the automatic approach that exploits both the work on MR-images done in Rome and the additional potential that lies in the use of automatic image analysis. For an automatic approach it is not necessarily useful to mimic the manual approach directly, as an automated analysis has both different strengths and different weaknesses from that of the manual process.

Our strategy will be to base the automatic MR-analysis on the development stages and bones as defined by the University of Rome, and then try to determine the development stages automatically from the images. We have chosen to go through development stages and not estimate bone age directly from the images, both because this is the way the manual procedure works and because we think this is a sensible strategy when several models for the correspondence between bone age and development might be needed. Direct estimation of bone age from the images would require more training data and models trained from the images for each group.

Within the limited resources of this study there was a need to limit the scope, and we have therefore studied only one of the bones in the hand that are treated in the manual process.

We have then chosen to study the radial bone, as this (together with the ulnar) is the bone that grows and develops through the whole age range. If we are able to demonstrate a method that works for the radial, this will also indicate a similar potential for the other bones of the hand.

In the analysis we have chosen to select a number of consecutive slices that are analyzed individually rather than constructing a 3D-model of the hand. We use more slices than that of the manual process, and select the slices to use based on simple strategies.

Furthermore, we have for the analysis of the bone used an approach that selects a rectangular region of interest (ROI) surrounding the bone end and fusion plate. This is similar to what is seen for more recent work on automatic analysis of X-rays and for MR-images, and this region also covers the areas considered in the manual analysis. The localization of ROIs is a much studied area for this problem, and many approaches have been suggested in the literature and may be used in a final system. Due to limited resources we have therefore chosen not to focus very much on this step, and just used a very simple approach which is fast to implement, and then just made manual adjustments of the results where needed.

From the ROI we will extract image-based features that are suitable for classifying the ROI into development stages. We have chosen to not base these features directly on the textual descriptions of the characteristics defined in the manual approach. Instead we will extract features that are more general, but that should be able to capture much of the same type of information. Based on these features we will do a classification into development stages.

The manual approach uses a score derived from the sum of development stages and a look-up-table to estimate the final bone age. In this limited study, where only the radial is considered, such an approach is not applicable. Furthermore, the use of a summed score may not be the best approach for an automatic solution. For our limited radial-based method, we have just defined a simple approach based on the age ranges defined for the development stages.

5.5 The automatic approach in more detail

Based on the assessment above, we have then focused the analysis in this study on a rectangular region covering the areas of interest for the radial bone, using a simple approach to localize this area. We localize this region from several consecutive slices of the MR sequence. The slices are selected as a sequence of 5 slices, centered around a fixed mid-slice found from the average of the slice numbers used in the manual approach. The MR-images are acquired from the left hand of the right-handed and from the right hand of the left-handed. To avoid having to determine which hand that is imaged in each sequence, we just mirror each slice about the y-axis and use both versions of the slice.

From the identified region in each slice we then extract a set of generic features. These features are found using methods based on so-called deep learning. These are deep neural nets that are able to learn data representations directly from the image data and that have obtained ground breaking results in recognition from complex images (Krizhevsky et al., 2012). In this study, we have had a very limited dataset which is much too small to train such a deep neural net. However, it has also been shown that these methods can learn generic features from an application area where a large training set is available, which can then be applied for another problem area where less data are available (Azizpour et al., 2014). Hence, we have used such a pre-trained net to extract the features from the region of interest surrounding the radial bone.

Based on these extracted features we do an automatic classification into development stage for each selected slice from the MR-sequence. This means that for each individual there will be a set of classifications corresponding to the number of selected slices. The classification results will not necessarily be the same for all slices. When one representative development stage is needed for each individual, we can then for instance select the most frequent class (development stage).

To be able to compare the results with studies that report estimated bone age compared to chronological age, we have also used the results to estimate bone age. As we in this study have only looked at the development of the radial bone, we can not use the approach defined by the University of Rome which is based on the sum of stages for 9 bones. Instead we have used an approach that estimates bone age as the average of the midpoints of the age intervals corresponding to the development stages that have been found for each slice.

5.6 Test and evaluation

The automatic approach has been evaluated on a set consisting of MR-sequences of 63 boys between 14 and 19 years. These are images of healthy Italian school boys that have been acquired by the University of Rome, using a small MR machine of 0.2 Tesla. We have tested our automatic approach on this dataset and compared the automatically derived development stages to the stages that have been manually determined for the radial bone by the researchers in Rome. We have also compared the automatically derived estimates of bone age with chronological age and compared the deviations from chronological age with that achieved with the manual approach and another automatic approach.

The results from our automatic classification into development stages, show that we get approximately 75% correspondence with the manually determined development stages for the radial bone. For the cases where the automatic classification results differ from the manual classifications, they never deviate with more than one stage in our experiments. Figure 17 shows these results as a confusion matrix, where the yellow entries along the diagonal are the classifications that correspond to the manually determined stages. The off-diagonal entries do not correspond exactly to the manually determined stages, but never deviate with more than one stage.

As the development stages are categorical (not continuous), the difference between the approaches in terms of derived stages will not necessarily capture the difference in accuracy of the bone age estimation. We have therefore also derived the bone age estimates using our automatic approach, and compared this to the chronological age for the individuals in the dataset. In Figure 17, the difference is reported in terms of mean deviation, absolute mean deviation, standard deviation and Root Mean Square error (RMS error). The same comparison has been done for the manually derived bone age estimates as received from the University of Rome. These estimates are based on all the nine bones as defined for their approach. Finally we have compared the absolute mean deviation and standard deviation to that reported for the automatic MRI-based approach from the University of Graz (Stern et al., 2014a). These have however been obtained on a different dataset and are not directly comparable.

	Manually derived develeopment stage					
	Stage	**8**	**9**	**10**	**11**	**12**
Automatically derived development stage	**8**	4				
	9	1	15	5		
	10		4	5	2	
	11			1	7	1
	12				2	16

Figure 17. Confusion matrix for the classification into development stages. The yellow entries along the diagonal are the classifications that correspond to the manually determined stages. The off-diagonal entries are the ones that do not correspond exactly to the manually determined stages, but they never deviate with more than one stage for this dataset.

Table 5 summarizes the results for the three approaches. As can be seen from this table, the manually and the automatically obtained results for the Italian dataset are very similar. Although not directly comparable, the deviations are also considerably lower than that obtained with the automatic approach from Graz.

Table 5. Deviation between chronological age and estimated bone age for the different approaches. *) The automatic approach from Graz has not been tested on the same dataset as the two other approaches, and in their study only the absolute mean deviation and standard deviation are reported.

Method	Mean deviation	Absolute mean deviation	Standard deviation	RMS error
Rome manual	-0.17	0.50	0.41	1.61
NR automatic	-0.14	0.47	0.42	1.58
Graz automatic*)	NA	0.85	0.58	NA

Figure 18 shows the deviations between predicted bone age and chronological age for the manual predictions (on the left) and for the automatic predictions on the Italian dataset (on the right).

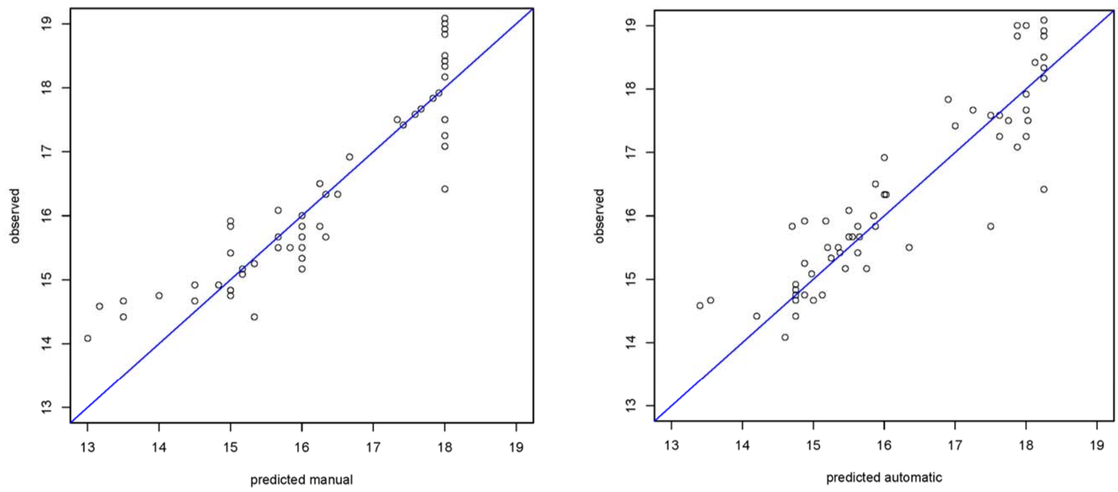


Figure 18. Plot of chronological age versus estimated bone age for the manual prediction from the University of Rome (left) and for the suggested automatic approach from the Norwegian Computing Center (right).

5.7 Conclusion

The conclusion so far is that for the current dataset, the automatic approach that we have suggested is able to give estimates with a precision very similar to that of the manually based MR method. This indicates that the suggested automatic approach has a similar potential for bone age estimation. However, both the automatic method and the manual approach need to be validated on a larger independent dataset before very firm conclusions can be drawn.

5.8 Recommendations

Use of MR images makes it possible to perform bone age estimation without the use of equipment generating ionizing radiation. This can again make it possible to acquire more up-to-date and larger reference datasets for future bone age estimation, which could be valuable. The manual approach developed at the University of Rome seems to be a viable approach for this. To verify this, it is however a need to validate the approach on a much larger and independent dataset.

The weakness of the manual approach is that it requires thorough analysis of many details in the whole MR-sequence, much more details than most X-ray based methods. Hence, this requires thorough training of the radiologists. Still, variability related to subjectivity and inter- and intra-observer differences can be a problem. Automation of the approach may therefore be even more desirable here than for X-ray based methods.

The results reported here using the suggested automatic approach based on MR images for bone age estimation are very promising. Hence, automation of the process should be possible. However, as for the manual approach, there is a need to validate the approach on a much larger independent dataset. Prior to further development of these methods, it is necessary to establish a large dataset of MR images from healthy individuals of known age. Such a large dataset does currently not exist, and an international effort is probably needed to collect such a set.

5.9 References

Azizpour H, Razavian AS, Sullivan J, Maki A, et al. (2014). From generic to specific deep representations for visual recognition. arXiv preprint, arXiv:1406.5774.

Ebner T, Stern D, Donner R, Bischof H, et al. (2014). Towards automatic bone age estimation from MRI: localization of 3D anatomical landmarks. Med Image Comput Comput Assist Interv., Vol. 17(Pt 2):421-428.

Eikvil L. (2016). Bruk av MR-bilder for estimering av skjelettalder. NR-Note, SAMBA/09/16.

Greulich WW, Pyle SI. (1959). Radiographic Atlas of Skeletal Development of the Hand and Wrist. Stanford, CA: Stanford University Press, 2nd edition.

Krizhevsky A, Sutskever I, Hinton GE. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems 25 (NIPS 2012), pp. 1097-1105.

Stern D, Ebner T, Bischof H, Grassegger S, et al. (2014). Fully automatic bone age estimation from left hand MR images. Med Image Comput Comput Assist Interv., Vol. 17(Pt 2):220-227.

Stern D, Ebner T, Bischof H, Urschler M. (2014). Determination of legal majority age from 3D magnetic resonance images of the radius bone. In 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI 2014), pp. 1119-1122.

Tanner JM, Whitehouse RH, Marshall WA, et al. (1975). Assessment of Skeletal Maturity and Prediction of Adult Height (TW2 Method). New York: Academic Press.

Serinelli S, Arbarello P, Battisti S, Tomei E, Semelka RC. (2014). Bone Age: Medico-legal Issues. Chapter 2 in Text-Atlas of Skeletal Age Determination: MRI of the Hand and Wrist in Children. ISBN: 978-1-118-69227-1. Wiley-Blackwell.

Tomei E, Sartori A, Nissman D, Al Ansari N, et al. (2014). Value of MRI of the hand and the wrist in evaluation of bone age: preliminary results. J Magn Reson Imaging: JMRI, Vol. 39(5):1198-1205.

Tomei E, Semelka RC, Nissman D. (2014). Text-Atlas of Skeletal Age Determination: MRI of the Hand and Wrist in Children. ISBN: 978-1-118-69227-1. Wiley-Blackwell.

Urschler M, Grassegger S, Stern D. (2015). What automated age estimation of hand and wrist MRI data tells us about skeletal maturation in male adolescents. Ann Hum Biol., Vol. 42(4):358-367.

6 WP4: Further development of new methods

Teeth and bones change shape/structure as they continue to grow until they reach full maturation in late adolescence or early adulthood. These changes are used to assess age. At present, the most commonly used methods for estimating age from teeth and bones assign a continuous process of maturation into discrete stages. WP4 has been a pilot study to explore methods that quantify dental maturation to improve reliability. The objective of this activity has been to establish a platform for further research in this field, by preparing and developing one or more PhD programs to collate these findings and establish a unified holistic approach to estimate age from teeth.

The following summarizes the work that has been done within WP4, both the literature review and the preparation and development of a PhD program.

Sigrid Ingeborg Kvaal has done a literature review and initial exploration of alternative dental approaches for quantifying maturation and development. A project description for a PhD program was prepared by Sigrid Ingeborg Kvaal and Co-Supervisor Professor Patrick Thevissen, Katholieke Univeristeit, Leuven, Belgium.

6.1 Background

Age assessment includes some form of staging or measurement of the development from childhood to an adult fully grown person. It measures growth and development such as height, weight, sexual maturity, dental and skeletal development and psychological maturations [1]. The recommendations are that age estimations are from more than one independent physical trait in the same individual but there exists no methods to combine different growth parameters. Skeletal and dental age developments are independent of each other [2] and are recommended used in age estimation procedures [3]. Existing reference datasets do only contain one type of measurement per individual (only wrist or only teeth). Furthermore, they are based on limited datasets from few populations, they have age categories rather than continuous age, and the description of associated uncertainties is incomplete.

Skeletal age estimations in children and young adults are commonly graded from the development and closure of the bony symphysis. The most widely used method is grading development of bones in the hand and wrist, but clavicle and ribs are also used [4-8]. Traditionally these comparisons have been made from analogue radiographs, but over the past decade new imaging techniques like computer tomography (CT), cone beam CT, microfocus CT, magnetic resonance imaging (MRI) and ultrasound are tested out.

The most commonly used method for age estimation from hand and wrist is based on the atlas of Greulich and Pyle [4]. Its latest edition was in 1959. It is based on 1.000 radiographs of children and contains reference images of left hand and wrist from children up to 18 years for females and 19 years for males and explanations regarding the gradual age related changes observed for each standard image. Bone age is calculated by comparing the left hand and wrist radiographs of the individual with the nearest matching reference radiograph provided in the atlas which are standard for different ages.

Gleiser and Hunt in 1955 first graded dental development from radiographs of the teeth [9]. This grading system into different stages was further developed by Moorrees and Fanning [10] and since then this grading of tooth development has been used with variations in staging from 10 to 14 stages [11-15]. Some investigation has only looked at the root development [16]. The principle of all these grading systems is that each stage of crown and root correspond to a mean or median tooth age. This applies to all 32 teeth and the age of the child is derived from the average of the mean or median tooth ages. Once the tooth is fully formed this tooth cannot be included. In her PhD thesis Haavikko [11] used this grading system and each tooth development was divided into 12 stages (6 for the crown and 6 for the root development) and the tooth age was when the tooth development was in the median of the stage. Other studies have used 10 stages [13,17]. Other studies from children in Canada and Scandinavia have also been advocated [18].

In young adolescents all teeth but the wisdom teeth have completed root development and only the staging of the four wisdom teeth can be used. Dental age estimations are made from the late stages of development of the roots on second and third molars (wisdom teeth) [19-21]. Frequently it is solely from wisdom teeth which are the only teeth not fully developed in this age group. In her work on development of the mandibular wisdom tooth from four different ethnic groups Liversidge used 14 stages (7 crown stages and 7 stages in root development) [14].

However, for dental age estimation in children the most commonly used method has been published by Demirijan et al. [22]. This is based on the same principle as Tanner and Whitehouse age estimation from bone development [23]. The grading of tooth development is based on 8 stages which in the original article were described in details and included pencil sketches and pictures of the radiographic stage. Each tooth is graded into stages of tooth development and labelled as A, B, C, D, ..., H which is then converted into a self weighted score. A dental maturity score is obtained by adding up the self weighted scores and this dental maturity score corresponds to an age which is different for males and females. The Demirijan method is based on French/Canadian children and includes teeth up to the age of 16 years. For adolescent children, Mincer, Harris and Berryman in 1993 [21] used Demirijan's 7 stages to grade the development of the wisdom tooth using separate tables for males and females and maxillary and mandibular teeth. This method to estimate age of maturity is widely used in America [24].

The critics of the staging system are that grading is subjective and even with training there will be some degree of inter-observer variations. Attempts have been made to make objective measurements by measuring tooth length and widths on radiographs [25]. Other studies have demonstrated that such measurements are not more accurate than subjective grading [19]. However recent studies seem to indicate the index between the width of open root apices and the height of the teeth is a good index for estimating ages older or younger than 18 years [26-28]. Due to biological variations exact age determination is with the present methods impossible, and only assessments can be made.

There is international agreement that the recommendation is to use two or more independent age estimation methods [3], but there are few publications which includes the combination of skeletal and dental age-related changes. Studies independent of dental or hand/wrist assessment have been carried out [29]. One study has combined age changes in the spine and dental development [7] and a few studies have compaired age estimation from dental development and hand and wrist changes in the same child [30-33]. In all reports the skeletal age is estimated using the Greulich and Pyle method [4]. The dental age estimation uses Demirijan's grading and estimates age according to methods described by Demirijan et al. [22] or Mincer et al. [21]. One study which includes children 5-15 years, do not compare the skeletal and dental age estimation methods but investigates which method estimates age closest to chronological age in different age groups [29]. Studies have also shown that the teeth show greater variation in adolescents [34].

6.2 Review of literature on 3D

Review of the literature was included as part of the PhD program (see section 6.3). CT imaging makes continuous sections of pictures which may be converted into 3D pictures. This technique has been used to estimate age from clavicular symphysis [8] and distal femur [35]. MRI present non-radiation images and has been used on iliac crest [36] and distal femur [35]. Only few studies have looked at the analysis of dental development from 3D pictures [37,38]. Age assessment has been attempted by measuring ratio between pulp and tooth volume in mandibular canines on CT with no significant difference between chronological age and estimated age. Sex assessment from total volume was correct in all cases [39]. An Australian study showed that conventional CT is not very accurate in estimating age from wisdom teeth, but can be used in forensic cases to discriminate between children and adults as 100% of females and 96% of males had fully developed wisdom teeth by the age of 18 years [38]. CTs of scull and upper body also give the advantage that two or more age-related changes can be included. These may develop independently and as such narrow the range in age estimations [40].

More recent studies related to age and 3D was made with microfocus CT [41-43]. In recent years cone beam CT has become popular among practising dentists and material has become available for studies. Studies have been carried out on all teeth [44] or only selected types of teeth [45,46]. Some studies have used the same grading system as on conventional radiographs [45] with correlation coefficients r^2=0.80 for males and r^2=0.78 for females. Several studies have calculated ratios between volume of pulp and hard dental tissue and showed that the volume ratios between pulp and tooth are related to age [44, 46-49].

MRI has also been investigated for use in dental age estimation but this is only in its early infancy [50,51]. Like CT images, dental MRI pictures can be used for the dental age assessment with equally good results as when using conventional radiographs [52,53]. Further analyses and improvement of methods are necessary before 3D image techniques can be used routinely in forensic age estimations.

6.3 PhD program

Here we give a brief description of the PhD program developed in this work package.

6.3.1 Aims

The aims of the PhD program were to use 3D pictures in radiology, which will give a more detailed image of tooth development, and to use continuous data, which will allow for more precise description of variation and uncertainty.

3D imaging techniques of interest are computer tomography (CT) and magnetic resonance imaging (MRI). At present, more material is available within CT so the focus in the first place was to develop techniques from CT images.

6.3.2 Material and methods

Material will be collected from routine computer tomographic from autopsies at Institute of Public Health, Division of Forensic Science, Oslo and Department of Forensic Medicine, University of Copenhagen.

Norwegian Computing Center (NR) has long traditions for image analysis and is at present working on projects with Faculty of Dentistry, University of Oslo.

6.3.3 Project plan

Suggested topics for the PhD program were:

- Literature review of 3D images used in age estimation.
- Data collection of medical imaging files, 3D and 2D.
- Develop automated variable recognition, segmentation and classification on 3D data.
- Validation of automated variable registration: (i) Compare with conventional 2D results and (ii) tests on reproducibility.

A candidate was interested in the project and an application was made to the Faculty of Dentistry for a PhD research position in February 2015. The candidate was not successful in the application.

6.4 References

[1] Eikvil L, Kvaal SI, Teigland A, Haugen M, Grøgaard J. (2012). Age estimation in youth and young adults: a summary of the needs for methodological research and development. NR-Note, SAMBA 52/12.

[2] Gelbrich B, Frerking C, Weiss S, Schwerdt S, et al. (2015). Combining wrist age and third molars in forensic age estimation: how to calculate the joint age estimate and its error rate in age diagnostics. Ann Hum Biol., Vol. 42(4):389-396.

[3] Schmeling A, Grundmann C, Fuhrmann A, Kaatsch HJ, et al. (2008). Criteria for age estimation in living individuals. Int J Legal Med., Vol. 122(6):457-460.

[4] Greulich WW, Pyle SI. (1959). Radiographic Atlas of Skeletal Development of the

Hand and Wrist. Stanford, CA: Stanford University Press, 2nd edition.

[5] Tanner JM, Whitehouse RH, Healy MJR. (1962). A new system for estimating skeletal maturity from the hand and wrist, with standards derived from a study of 2600 healthy British children. Part II: The scoring system. Paris: International Children's Centre.

[6] Schulz R, Muhler M, Reisinger W, Schmidt S, et al. (2008). Radiographic staging of ossification of the medial clavicular epiphysis. Int J Legal Med., Vol. 122(1):55-58.

[7] Thevissen PW, Kaur J, Willems G. (2012). Human age estimation combining third molar and skeletal development. Int J Legal Med., Vol. 126(2):285-292.

[8] Kellinghaus M, Schulz R, Vieth V, Schmidt S, et al. (2009). Forensic age estimation in living subjects based on the ossification status of the medical clavicular epiphysis as revealed by thin slice multidetector computed tomography. Int J Legal Med., Vol. 124:149-154.

[9] Gleiser I, Hunt EE Jr. (1955). The permanent mandibular first molar: its calcification, eruption and decay. Am J Phys Anthropol., Vol. 13(2):253-283.

[10] Moorrees CFA, Fanning EA, Hunt EE Jr. (1963). Age variation of formation stages for ten permanent teeth. J Dent Res., Vol. 42(6):1490-1502.

[11] Haavikko K. (1970). The formation and the alveolar and clinical eruption of the permanent teeth. An orthopantomographic study. PhD thesis, University of Helsinki, Finland.

[12] Liversidge HM. (2012). The assessment and interpretation of Demirjian, Goldstein and Tanner's dental maturity. Ann Hum Biol., Vol. 39(5):412-431.

[13] Nolla CM. (1960). The development of the permanent teeth. J Dent Child., 254-266.

[14] Liversidge HM. (2008). Timing of human mandibular third molar formation. Ann Hum Biol., Vol. 35(3):294-321.

[15] AlQahtani SJ, Hector MP, Liversidge HM. (2010). Brief communication: The London atlas of human tooth development and eruption. Am J Phys Anthropol., Vol. 142(3):481-490.

[16] Kullman L, Johanson G, Akesson L. (1992). Root development of the lower third molar and its relation to chronological age. Swed Dent J., Vol. 16(4):161-167.

[17] Kohler S, Schmelzle R, Loitz C, Puschel K. (1994). Die Entwicklung des Weisheitszahnes als Kriterium der Lebensaltersbestimmung. Ann Anat., Vol. 176(4):339-345.

[18] Anderson DL, Thompson GW, Popovich F. (1976). Age of attainment of mineralization stages of the permanent dentition. J Forensic Sci., Vol. 21(1):191-200.

[19] Thevissen PW, Fieuws S, Willems G. (2013). Third molar development: evaluation of nine tooth development registration techniques for age estimations. J Forensic Sci., Vol. 58(2):393-397.

[20] Willems G, Van Olmen A, Spiessens B, Carels C. (2001). Dental age estimation in Belgian children: Demirjian's technique revisited. J Forensic Sci., Vol 46(4):893-895.

[21] Mincer HH, Harris EF, Berryman HE. (1993). The A.B.F.O. study of third molar development and its use as an estimator of chronological age. J Forensic Sci., Vol. 38:379-90.

[22] Demirjian A, Goldstein H, Tanner JM. (1973). A new system of dental age assessment. Hum Biol., Vol. 45(2):211-227.

[23] Demirjian A. (1978). Dentition. In: Falkner F, Tanner JM (eds.), Human Growth - 2 Postnatal Growth, pp. 413-444. Springer US.

[24] Lewis JM, Senn DR. (2010). Dental age estimation utilizing third molar development: A review of principles, methods, and population studies used in the United States. Forensic Sci Int., Vol. 201(1-3):79-83.

[25] Mörnstad H, Staaf V, Welander U. (1994). Age estimation with the aid of tooth development: a new method based on objective measurements. Scand J Dent Res., Vol. 102(3):137-143.

[26] Cameriere R, Ferrante L, De Angelis D, Scarpino F, et al. (2008). The comparison between measurement of open apices of third molars and Demirjian stages to test chronological age of over 18 year olds in living subjects. Int J Legal Med., Vol. 122(6):493-497.

[27] De Luca S, Biagi R, Begnoni G, Farronato G, et al. (2014). Accuracy of Cameriere's cut-off value for third molar in assessing 18 years of age. Forensic Sci Int., Vol. 235:102.e1-6.

[28] Cameriere R, Santoro V, Roca R, Lozito P, et al. (2014). Assessment of legal adult age of 18 by measurement of open apices of the third molars: Study on the Albanian sample. Forensic Sci Int., Vol. 245C:205.e1-205.e5.

[29] Pechnikova M, Gibelli D, De Angelis D, de Santis F, et al. (2011). The "blind age assessment": applicability of Greulich and Pyle, Demirjian and Mincer aging methods to a population of unknown ethnic origin. Radiol Med., Vol. 116(7):1105-1114.

[30] Suma GN, Rao BB, Annigeri RG, Rao DJK, et al. (2011). Radiographic correlation of dental and skeletal age: Third molar, an age indicator. J Forensic Dent Sci., Vol. 3(1):14-18.

[31] Varkkola O, Ranta H, Metsäniitty M, Sajantila A. (2011). Age assessment by the Greulich and Pyle method compared to other skeletal X-ray and dental methods in data from Finnish child victims of the Southeast Asian Tsunami. Forensic Sci Med Pathol., Vol. 7(4):311-316.

[32] Shilpa PH, Sunil RS, Sapna K, Kumar NC. (2013). Estimation and comparison of dental, skeletal and chronologic age in Bangalore south school going children. J Indian Soc Pedod Prev Dent., Vol. 31(2):63-68.

[33] Santoro V, Roca R, De Donno A, Fiandaca C, et al. (2012). Applicability of Greulich and Pyle and Demirijan aging methods to a sample of Italian population. Forensic Sci Int., Vol. 221(1-3):153.e1-5.

[34] Kullman L. (1995). Accuracy of two dental and one skeletal age estimation method in Swedish adolescents. Forensic Sci Int., Vol. 75(2-3):225-236.

[35] Krämer JA, Schmidt S, Jürgens KU, Lentschig M, et al. (2014). Forensic age estimation in living individuals using 3.0 T MRI of the distal femur. Int J Legal Med., Vol. 128(3):509-514.

[36] Wittschieber D, Vieth V, Timme M, Dvorak J, et al. (2014). Magnetic resonance imaging of the iliac crest: age estimation in under-20 soccer players. Forensic Sci Med Pathol., Vol. 10(2):198-202.

[37] Cantekin K, Ercan Sekerci A, Peduk K, Delikan E, et al. (2014). Dental age assessment for different climatic regions. Am J Forensic Med Pathol., Vol. 35(3):197-200.

[38] Bassed RB, Briggs C, Drummer OH. (2011). Age estimation and the developing third molar tooth: an analysis of an Australian population using computed tomography. J Forensic Sci., Vol. 56(5):1185-1191.

[39] Tardivo D, Sastre J, Ruquet M, Thollon L, et al. (2011). Three-dimensional modeling of the various volumes of canines to determine age and sex: a preliminary study. J Forensic Sci., Vol. 56(3):766-770.

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

[41] Aboshi H, Takahashi T, Komuro T. (2010). Age estimation using microfocus X-ray computed tomography of lower premolars. Forensic Sci Int., Vol.200(1-3):35-40.

[42] Someda H, Saka H, Matsunaga S, Ide Y, et al. (2009). Age estimation based on three-dimensional measurement of mandibular central incisors in Japanese. Forensic Sci Int., Vol. 185(1-3):110-114.

[43] Vandevoort FM, Bergmans L, Van Cleynenbreugel J, Bielen DJ, et al. (2004). Age calculation using X-ray microfocus computed tomographical scanning of teeth: a pilot study. J Forensic Sci., Vol. 49(4):787-790.

[44] Maret D, Peters OA, Dedouit F, Telmon N, et al. (2011). Cone-Beam Computed Tomography: a useful tool for dental age estimation? Med Hypotheses., Vol. 76(5):700-702.

[45] Cantekin K, Sekerci AE, Buyuk SK. (2013). Dental computed tomographic imaging as age estimation: morphological analysis of the third molar of a group of Turkish population. Am J Forensic Med Pathol., Vol. 34(4):357-362.

[46] Ge ZP, Ma RH, Li G, Zhang JZ, et al. (2015). Age estimation based on pulp chamber volume of first molars from cone-beam computed tomography images. Forensic Sci

Int., Vol. 253:133.e1-7.

[47] Pinchi V, Pradella F, Buti J, Baldinotti C, et al. (2015). A new age estimation procedure based on the 3D CBCT study of the pulp cavity and hard tissues of the teeth for forensic purposes: A pilot study. J Forensic Leg Med., Vol. 36:150-157.

[48] Star H, Thevissen P, Jacobs R, Fieuws S, et al. (2011). Human dental age estimation by calculation of pulp-tooth volume ratios yielded on clinically acquired cone beam computed tomography images of monoradicular teeth. J Forensic Sci., Vol. 56 Suppl 1:S77-82.

[49] Yang F, Jacobs R, Willems G. (2006). Dental age estimation through volume matching of teeth imaged by cone-beam CT. Forensic Sci Int., Vol. 159 Suppl 1:S78-83.

[50] Hövener JB, Zwick S, Leupold J, Eisenbeibeta AK, et al. (2012). Dental MRI: imaging of soft and solid components without ionizing radiation. J Magn Reson Imaging., Vol. 36(4):841-846.

[51] Rasche V, Hofmann C, Maschka R, Ulrici J, et al. (2009). The application of ultr-short echo time MRI (UTE) for the structural assessment of dental hard-tissue components. IFMBE proceedings, Vol. 25/11:65-67.

[52] Baumann P, Widek T, Merkens H, Boldt J, et al. (2015). Dental age estimation of living persons: Comparison of MRI with OPG. Forensic Sci Int., Vol. 253:76-80.

[53] Guo Y, Olze A, Ottow C, Schmidt S, et al. (2015). Dental age estimation in living individuals using 3.0 T MRI of lower third molars. Int J Legal Med., Vol. 129(6):1265-1270.