

Sammenligning av manuell og automatisert skjelettvurdering



Foto: iStockphoto.com

Notatnr
Forfatter

SAMBA/13/16
Marion Haugen
Jens B. Grøgaard
Tor Ole Kjellevand

Dato

26. april 2016

Norsk Regnesentral

Norsk Regnesentral (NR) er en privat, uavhengig stiftelse som utfører oppdragsforskning for bedrifter og det offentlige i det norske og internasjonale markedet. NR ble etablert i 1952 og har kontorer i Kristen Nygaards hus ved Universitetet i Oslo. NR er et av Europas største miljøer innen anvendt statistisk-matematisk modellering og har et senter for forskningsdrevet innovasjon, Big Insight, med finansiering fra Norges forskningsråd, bedrifter og offentlige partnere. Innen statistikk jobbes det med et bredt spekter av problemstillinger, for eksempel finansiell risiko, jordobservasjon, estimering av fiskebestander, helse og beskrivelse av geologien i petroleumsreservoarer. NR er ledende i Norge innen utvalgte deler av informasjons- og kommunikasjonsteknologi. Innen IKT-området har NR innsatsområdene e-inkludering, informasjonssikkerhet og smarte informasjonssystemer.

NRs visjon er forskningsresultater som brukes og synes.

Tittel	Sammenligning av manuell og automatisert skjelettvurdering
Forfatter	Marion Haugen <Marion.Haugen@nr.no> Jens B. Grøgaard <Jens.Grogaard@helsedir.no> Tor Ole Kjellevand <Tor.Ole.Kjellevand@unilabs.com>
Dato	26. april 2016
Publikasjonsnummer	SAMBA/13/16

Sammendrag

I Norge benyttes en medisinsk aldersvurdering dersom det ikke er mulig å fastsette med rimelig sikkerhet hvor gammel en enslig mindreårig asylsøker er. En røntgenundersøkelse av hånd og håndrot, der benveksten undersøkes manuelt, er den ene av to deler av en medisinsk aldersvurdering. Den visuelle analysen utføres av en radiolog ved at røntgenbildet som skal analyseres blir manuelt sammenlignet med bildeatlasen til Greulich og Pyle for å finne referansen som ligner mest, og ut i fra det bestemmes skjelettalder.

BoneXpert er programvare for automatisert estimering av skjelettalder fra røntgenbilder. Programvaren baserer seg på bildeanalyse av røntgenbilder av hånd og håndrot. Hensikten med automatisering er å gjøre skjelettvurderingen mer objektiv og mindre avhengig av radiologene som utfører de visuelle analysene.

I dette notatet har vi testet BoneXpert på 100 utvalgte røntgenbilder av gutter som det tidligere har blitt utført en manuell skjelettvurdering på. Vi har sammenlignet skjelettalder fra manuell og automatisert skjelettvurdering.

Testen viser at det er over 90% samsvar mellom manuell og automatisert skjelettvurdering, i henhold til om skjelettalder er over eller under 18 år. Det gode samsvaret 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.

Emneord	Skjelettalder, enslige mindreårige asylsøkere, manuell, automatisert, samsvar, avvik
Målgruppe	UDI, aktører involvert i aldersvurderinger i Norge
Tilgjengelighet	Åpen
Prosjekt	FOUALDER
Prosjektnummer	220642
Satsningsområde	Statistikk for teknologi, industri og forvaltning
Antall sider	24
© Copyright	Norsk Regnesentral

Innhold

Liste med figurer	7
Liste med tabeller	7
English summary	9
1 Innledning	11
2 Data og metoder	13
2.1 Manuell skjelettvurdering	14
2.2 Automatisert skjelettvurdering: BoneXpert	14
2.3 Samsvar mellom de to skjelettvurderingene	15
2.4 Avvik mellom de to skjelettvurderingene	16
3 Resultater	17
3.1 Kriterium 1	19
3.2 Kriterium 2	20
4 Oppsummering og konklusjon	23
Referanser	24

Figurer

1	Resultatet fra den endelige medisinske aldersvurderingen utført på 100 utvalgte gutter	13
2	Eksempler på røntgenbilder av venstre hånd og håndrot	15
3	Spredningsplott av automatisert skjelettvurdering mot manuell skjelettvurdering utført på 100 utvalgte gutter	17
4	Manuell og automatisert skjelettvurdering utført på 100 utvalgte gutter . .	18
5	Prosentandel observasjoner med og uten samsvar fra aldersvurderinger av 100 utvalgte gutter (kriterium 1)	19
6	Prosentandel observasjoner med og uten samsvar fra aldersvurderinger av 100 utvalgte gutter (kriterium 2)	21

Tabeller

1	Avvik mellom manuell og automatisert skjelettvurdering for kriterium 1 .	20
2	Avvik mellom manuell og automatisert skjelettvurdering for kriterium 2 .	21

English summary

Introduction

Unaccompanied minor asylum seekers in Norway may be unaware of their exact date of birth, travel without appropriate identity documents or bring documents of uncertain provenance. For Norwegian authorities it is important to determine whether the asylum seeker is older or younger than 18 years of age. When this is impossible with reasonable certainty, a medical age assessment is performed.

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. Dental development is evaluated with a dental clinical examination and orthopantomogram (OPG) radiographs of the jaw at Institute of Clinical Dentistry, Faculty of Dentistry, University of Oslo. The private radiology institute Unilabs Norway evaluates the skeletal maturity by manually comparing radiographs of the left hand and wrist to the atlas of Greulich and Pyle (1959). The two independent age assessments are summarized in a final age assessment at BarneSak AS.

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; variability related to subjectivity, inter- and intra-observer differences. The purpose of automation is to make the skeletal assessment more objective and less dependent on the radiologists performing the visual analysis.

Unilabs Norway has tested BoneXpert on a sample of 100 left hand and wrist radiographs of boys. A manual skeletal age assessment of these boys has been performed by Unilabs Norway from August 2013 to December 2014. The purpose of this study has been to understand agreement and differences between the two skeletal age assessments.

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. To grade the development of bones in the left hand and wrist, a radiologist manually compares the radiographs of the individual to the atlas. The skeletal age is estimated from the nearest matching reference radiograph in the atlas. The task for Unilabs Norway is to give the skeletal age as an integer. According to the Greulich and Pyle atlas, 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 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 og 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: (i) The individual is 18 years or older from both of the skeletal age assessments (Agreement 1), (ii) the individual is below 18 years from both of the skeletal age assessments (Agreement 2), (iii) the individual is 18 years or older from the manual skeletal age assessment and below 18 years from the automated skeletal age assessment (Mismatch 1) and (iv) the individual is below 18 years from the manual skeletal age assessment and 18 years or older from the automated skeletal age assessment (Mismatch 2).

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. The skeletal age from BoneXpert must be 18 years or higher, or 17.5 years or higher, for the individual to be above 18 years. The second criterion is perhaps better suited to a continuous scale.

Results and conclusions

We have performed a successful test of BoneXpert on a sample of 100 left hand and wrist radiographs 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.

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, there is a need for a larger study. 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%. With a larger sample of radiographs, a percentage error related to the operational standard can be calculated.

1 Innledning

Dersom enslige mindreårige asylsøkere (EMA)¹ ikke kjenner til egen alder eller det er tvil rundt oppgitt alder benyttes metoder for aldersvurdering i Norge. Den medisinske aldersvurderingen består av to deler som er biologisk uavhengige (Gelbrich et al., 2015). Den ene delen er en tannundersøkelse som inkluderer en klinisk undersøkelse av tenner og munnhule og et stort røntgenbilde (OPG) som inkluderer alle tennene, tannrøttene og kjevene for å gradere krone- og rotutvikling. Den andre delen er en røntgenundersøkelse av hånd og håndrot der benveksten undersøkes. Det er skjelettvurderingen av hånd og håndrot vi skal se på i dette notatet.

De to delene i den medisinske aldersvurderingen utføres på to adskilte institusjoner, som ikke kjenner til hva den andre gjør eller konkluderer med. Institutt for klinisk odontologi, Det odontologiske fakultet, Universitetet i Oslo er ansvarlig for aldersvurderinger basert på tannutvikling. Det private røntgeninstituttet Unilabs Norge har vært ansvarlig for den manuelle skjelettvurderingen siden 1. august 2013. En barneradiolog vurderer asylsøkerens alder på bakgrunn av skjelettmodningen. Fra 1. august 2013 har BarneSak AS sammenstilt resultatene av disse to uavhengige undersøkelsene til en endelig medisinsk aldersvurdering og denne konklusjonen kalles samlet vurdering i dette notatet. Aldersestimatet basert på skjelettvurdering kalles også for skjelettalder i dette notatet.

Unilabs Norge har testet en metode for automatisert skjelettvurdering, BoneXpert, på et utvalg røntgenbilder av gutter som Unilabs Norge i perioden fra august 2013 til desember 2014 har utført en manuell skjelettvurdering på. Hensikten er å se på samsvar og avvik mellom skjelettalder fra manuell skjelettvurdering og skjelettalder fra automatisert skjelettvurdering.

Haugen et al. (2016) har sett på samsvar mellom manuell skjelettvurdering og tannvurdering på alle EMA som det har blitt utført en aldersvurdering av i Norge fra 2010 til 2014. Samsvar har de definert i henhold til om asylsøkers alder vurderes til å være over eller under 18 år. De har vurdert det til å være godt samsvar mellom de to uavhengige aldersestimatene basert på skjelett- og tannvurdering. I perioden fra august 2013 til desember 2014 er det utført 1.209 aldersvurderinger av gutter og det er 77.9% samsvar mellom manuell skjelettvurdering og tannvurdering. Det er utført 194 aldersvurderinger av jenter og det er 79.4% samsvar mellom manuell skjelettvurdering og tannvurdering.

Avsnitt 2 beskriver utvalget som er med i testen, metodene for manuell og automatisert skjelettvurdering, og hva vi mener med samsvar og avvik mellom de to skjelettvurderingene. Resultatene fra testen er beskrevet i avsnitt 3. Avsnitt 4 gir en kort oppsummering og konklusjon.

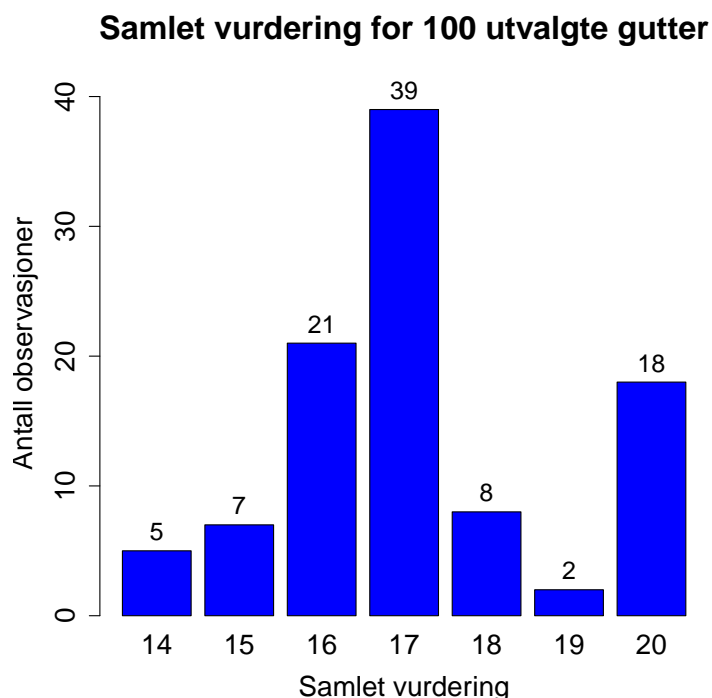
1. Enslig mindreårig asylsøker (EMA) er en asylsøker som oppgir å være under 18 år, og som ikke har følge av foreldre eller andre med foreldreansvar.

2 Data og metoder

Datasettet består av et utvalg på 100 gutter (EMA) som det har blitt utført en aldersvurdering av i Norge fra august 2013 til desember 2014. Det ble utført til sammen 1.209 aldersvurderinger på gutter i denne perioden (Haugen et al., 2016). Den manuelle skjelettvurderingen ble utført av det private røntgeninstituttet Unilabs Norge og BarneSak AS gjorde den endelige medisinske aldersvurderingen (samlet vurdering). Unilabs Norge har testet BoneXpert, et system for automatisert estimering av skjelettalder for røntgenbilder, på de 100 utvalgte guttene.

For den statistiske analysen brukes skjelettalder fra manuell skjelettvurdering (gitt som heltall), skjelettalder fra automatisert skjelettvurdering (gitt med to desimaler pga. kontinuerlig skala) og samlet vurdering (gitt som heltall).

For å sikre at vi fikk relevante individer med i testen, var det et krav til utvalget at de fleste guttene skulle ha alder over 15 år, gjerne flest over 17 år. Det er resultatet av den endelige medisinske aldersvurderingen (samlet vurdering) vi har brukt for å se på aldersfordelingen for gutter. Figur 1 viser samlet vurdering for de 100 guttene i utvalget, inndelt i syv aldersgrupper fra 14 år til 20 år. Det er 95 gutter som har samlet vurdering på 15 år eller høyere og 67 gutter som har samlet vurdering på 17 år eller høyere.



Figur 1. Resultatet fra den endelige medisinske aldersvurderingen (samlet vurdering) utført på 100 utvalgte gutter.

2.1 Manuell skjelettvurdering

Den manuelle skjelettvurderingen fra røntgenbilde av venstre hånd og håndrot er basert på metoden til Greulich og Pyle (1959). De har definert et bildeatlas som beskriver graden av skjelettutvikling knyttet opp mot alderstrinn, og atlasen er beregnet på visuell sammenligning. Manuell skjelettvurdering utføres ved at en barneradiolog sammenligner røntgenbildet av venstre hånd og håndrot med bildeatlasen til Greulich og Pyle for å finne referansen som ligner mest, og ut i fra det bestemme skjelettalder. Til venstre i figur 2 på side 15 vises et eksempel på et røntgenbilde av venstre hånd og håndrot.

Oppdraget til Unilabs Norge er å gi skjelettalderen i hele år. Skjelettet er vanligvis ferdig utviklet når gutter er 19 år, og da kan ikke en skjelettvurdering si noe annet enn at en gutt er 19 år eller eldre. Når skjelettet er ferdig utviklet settes derfor skjelettalder fra manuell vurdering til 19 år for gutter.

En klar ulempe med manuelle metoder er at det kan bli variasjoner knyttet til subjektivitet, og det kan være inter- og intra-observatør forskjeller.

2.2 Automatisert skjelettvurdering: BoneXpert

BoneXpert er programvare for automatisert estimering av skjelettalder fra røntgenbilder. Hensikten med automatisering er å gjøre aldersvurderingen mer objektiv og reproducerbar, og mindre avhengig av radiologene som utfører analysene. BoneXpert er utviklet i Danmark og baserer seg på bildeanalyse av røntgenbilder av hånd og håndrot. Programvaren tar utgangspunkt i metoder utviklet av administrerende direktør i Visiana, Hans Henrik Thodberg, med flere, og er beskrevet i Thodberg (2009) og Thodberg et al. (2009).

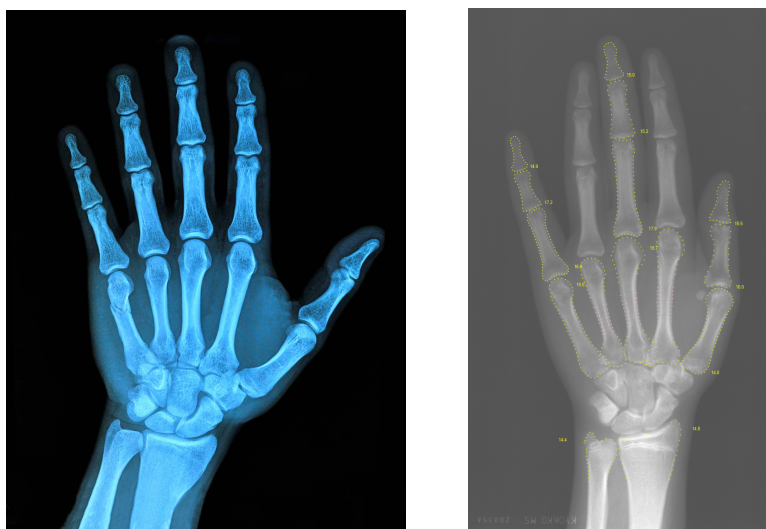
Den automatiserte skjelettvurderingen består av følgende tre hovedtrinn:

- (i) Omrisset av 13 bein i hånden finnes automatisk og konturer med feil forkastes.
- (ii) Egenskaper trekkes ut og det bestemmes en initiell skjelettalder for hvert bein. Potensielle feil forkastes.
- (iii) Resultatet fra forrige trinn transformeres til en skala som er kjent for radiologene. Det er skjelettalder fra Greulich og Pyle (1959) som er mest brukt, og som er brukt i vår test av programvaren.

Unilabs Norge har testet BoneXpert på 100 utvalgte røntgenbilder av gutter som de har utført en manuell skjelettvurdering på i perioden fra august 2013 til desember 2014. Programvaren som Unilabs Norge har testet er utviklet med basis i 231 røntgenbilder av friske sveitsiske barn født mellom 1954 og 1956. Skala for Greulich og Pyle skjelettalder for gutter er fra 2.5 år og opp til 19 år. Metoden er tilgjengelig som en prototype med versjon nummer 2.1.1.1. BoneXpert har en kontinuerlig skala og skjelettalderen er derfor gitt med to desimaler.

Norsk Regnesentral (NR) har testet BoneXpert på noen få røntgenbilder for å se hvordan programvaren fungerer, og til høyre i figur 2 på side 15 vises et eksempel på et røntgenbilde av venstre hånd og håndrot som NR har analysert med BoneXpert.

BoneXpert er godt testet ut i flere store studier for Greulich og Pyle skjelettalder 2.5-17 år for gutter og 2-15 år for jenter (Martin et al., 2013; Thodberg og Sävendahl, 2010; van Rijn et al., 2009). For gutter med skjelettalder opp til 17 år, kan alle korte bein i hånden brukes til skjelettvurderingen og skjelettalderen kan bestemmes som et gjennomsnitt av skjelettalder for ulike bein i hånden. Dette resulterer i en pålitelig skjelettvurdering. For gutter med skjelettalder høyere enn 17 år så er det kun beinene radius og ulna som fortsatt er under utvikling og som kan brukes til skjelettvurdering. Ulnabeinet er vanligvis ferdig utviklet når gutter er 18 år og skjelettvurderingen er da kun basert på radiusbeinet. Dermed er pålitelig lokalisering og tolkning av radius helt avgjørende for programvaren for gutter som er eldre enn 18 år.



Figur 2. Til venstre: Eksempel på et røntgenbilde av venstre hånd og håndrot (foto: iStockphoto.com). Til høyre: Eksempel på et røntgenbilde av venstre hånd og håndrot som Norsk Regnesentral har analysert med BoneXpert.

2.3 Samsvar mellom de to skjelettvurderingene

På samme måte som i Haugen et al. (2016) har vi definert samsvar mellom manuell og automatisert skjelettvurdering til enten å være at begge skjelettvurderingene gir en alder på 18 år eller eldre, eller at begge skjelettvurderingene gir en alder under 18 år. Det er fire mulige utfall for de to skjelettvurderingene, to utfall med samsvar og to utfall uten samsvar:

- **Samsvar 1:** Individet er 18 år eller eldre fra begge skjelettvurderingene.
- **Samsvar 2:** Individet er under 18 år fra begge skjelettvurderingene.
- **Ikke samsvar 1:** Individet er 18 år eller eldre fra manuell skjelettvurdering og under 18 år fra automatisert skjelettvurdering.
- **Ikke samsvar 2:** Individet er under 18 år fra manuell skjelettvurdering og 18 år eller eldre fra automatisert skjelettvurdering.

Det finnes ingen operativ standard for å avgjøre om individet er over eller under 18 år. Oppdraget i den manuelle skjelettvurderingen er å gi skjelettalderen i hele år og individet er dermed 18 år eller eldre dersom skjelettalderen er 18 år eller høyere. Automatisert skjelettvurdering gir kontinuerlig skjelettalder som gir mer fleksibilitet i valg av kriterium for å avgjøre om individet er over eller under 18 år.

Vi har sett på to ulike kriterier for å avgjøre om individet er over eller under 18 år basert på den automatiserte skjelettvurderingen:

- **Kriterium 1:** Skjelettalder fra automatisert skjelettvurdering må være 18 år eller høyere for at individet skal være 18 år eller eldre.
- **Kriterium 2:** Skjelettalder fra automatisert skjelettvurdering må være 17.5 år eller høyere for at individet skal være 18 år eller eldre.

Det første kriteriet er likt som den manuelle skjelettvurderingen. Det andre kriteriet er et naturlig kriterium i henhold til avrunding av et desimaltall til nærmeste hele tall og er kanskje bedre tilpasset en kontinuerlig skala.

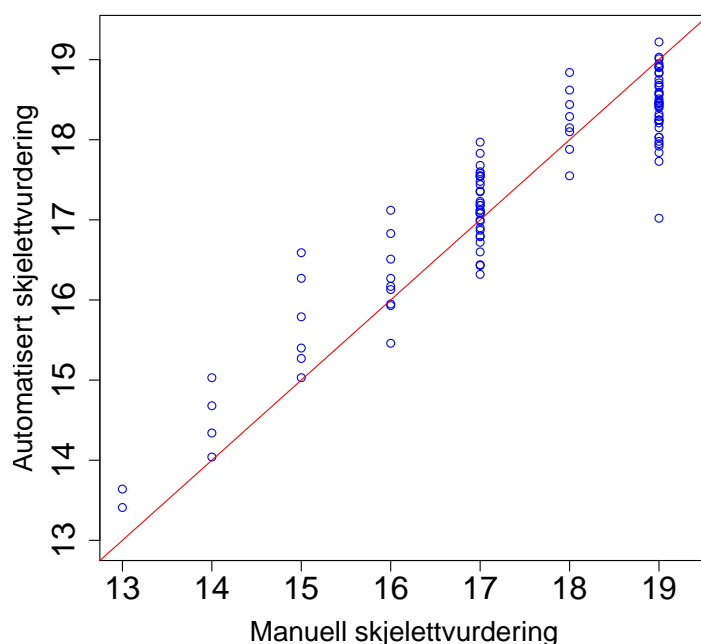
Det finnes ikke et objektivt mål på godt samsvar. Vi har derfor gjort en skjønnsvurdering av samsvaret vi har funnet mellom de to skjelettvurderingene.

2.4 Avvik mellom de to skjelettvurderingene

Avvik mellom de to skjelettvurderingene er definert som differansen mellom skjelettalder fra manuell skjelettvurdering og skjelettalder fra automatisert skjelettvurdering. Positive avvik betyr at individet er 18 år eller eldre fra manuell skjelettvurdering og under 18 år fra automatisert skjelettvurdering (Ikke samsvar 1). For negative avvik så er individet under 18 år fra manuell skjelettvurdering og 18 år eller eldre fra automatisert skjelettvurdering (Ikke samsvar 2).

3 Resultater

Figur 3 viser et spredningsplott av skjelettalder fra automatisert skjelettvurdering mot skjelettalder fra manuell skjelettvurdering for de 100 utvalgte guttene, sammen med linjen som viser om de to skjelettvurderingene er identiske (i rødt). Manuell skjelettvurdering er gitt som heltall og er inndelt i syv aldersgrupper fra 13 år til 19 år. Automatisert skjelettvurdering fra BoneXpert er gitt med to desimaler. Figuren viser at når manuell skjelettvurdering er i aldersgruppene fra 13 år til 16 år, så ser automatisert skjelettvurdering ut til å være høyere. For aldersgruppen 19 år så ser automatisert skjelettvurdering ut til å være lavere enn manuell skjelettvurdering.



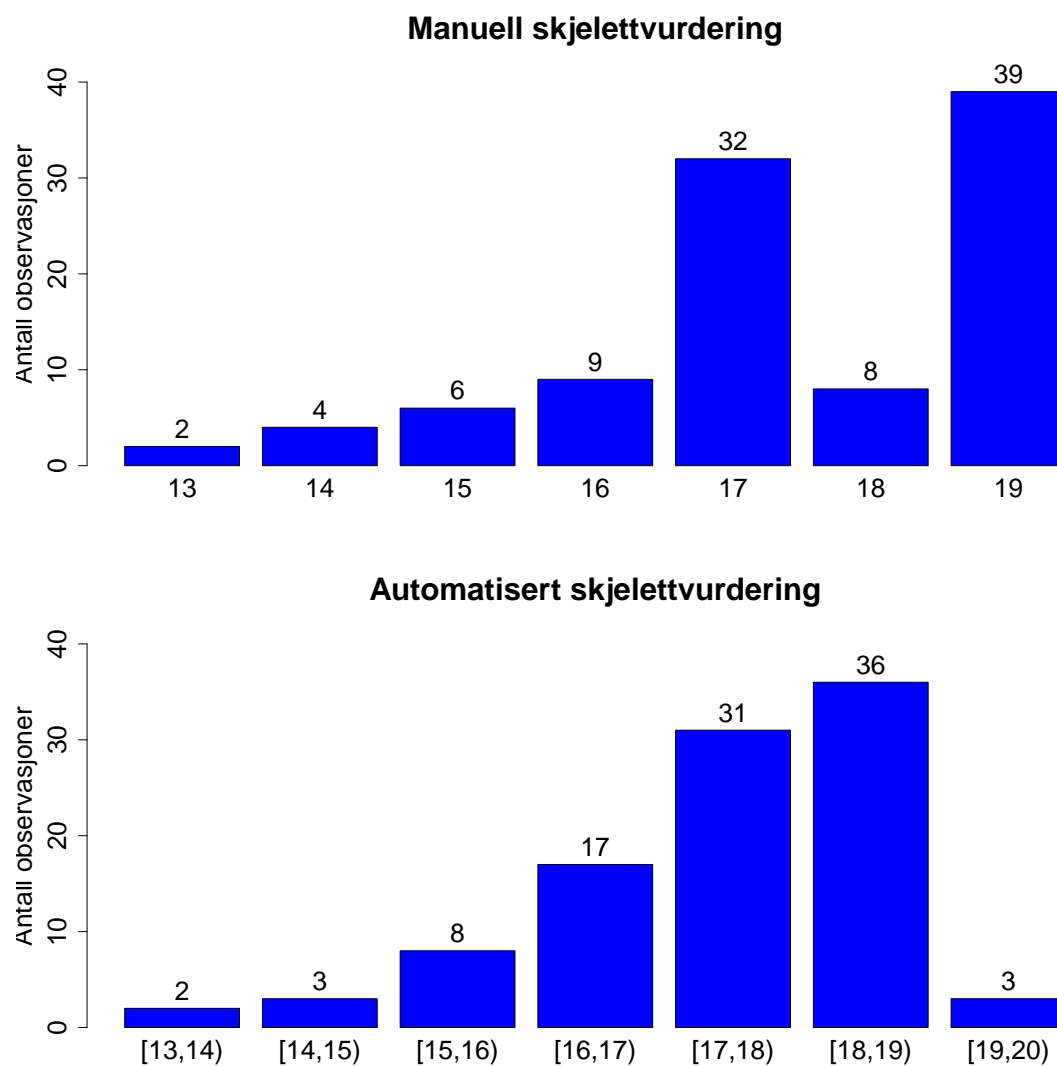
Figur 3. Spredningsplott av automatisert skjelettvurdering mot manuell skjelettvurdering utført på 100 utvalgte gutter. Identiske skjelettvurderinger ligger langs den røde linjen. Manuell skjelettvurdering er gitt som heltall og automatisert skjelettvurdering fra BoneXpert er gitt med to desimaler.

Root-Mean-Square Error (RMSE) er et mål på avviket mellom manuell og automatisert skjelettvurdering. La MS_i være manuell skjelettvurdering for observasjon nummer i og la AS_i være automatisert skjelettvurdering for observasjon nummer i , der $i = 1, \dots, n$ og $n = 100$ er antall observasjoner. Da beregnes RMSE slik:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (MS_i - AS_i)^2}{n}}$$

RMSE mellom manuell og automatisert skjelettvurdering i figur 3 er 0.61 år.

Figur 4 viser manuell (øverst) og automatisert (nederst) skjelettvurdering utført på 100 utvalgte gutter. Automatisert skjelettvurdering fra BoneXpert er gitt med to desimaler så aldersgruppene er gitt som syv halvåpne intervaller. Et halvåpent intervall $[a,b)$ omfatter alle reelle tall som er større enn eller lik a og mindre enn b . Aldersgruppene er fra $[13,14)$ til $[19,20)$.

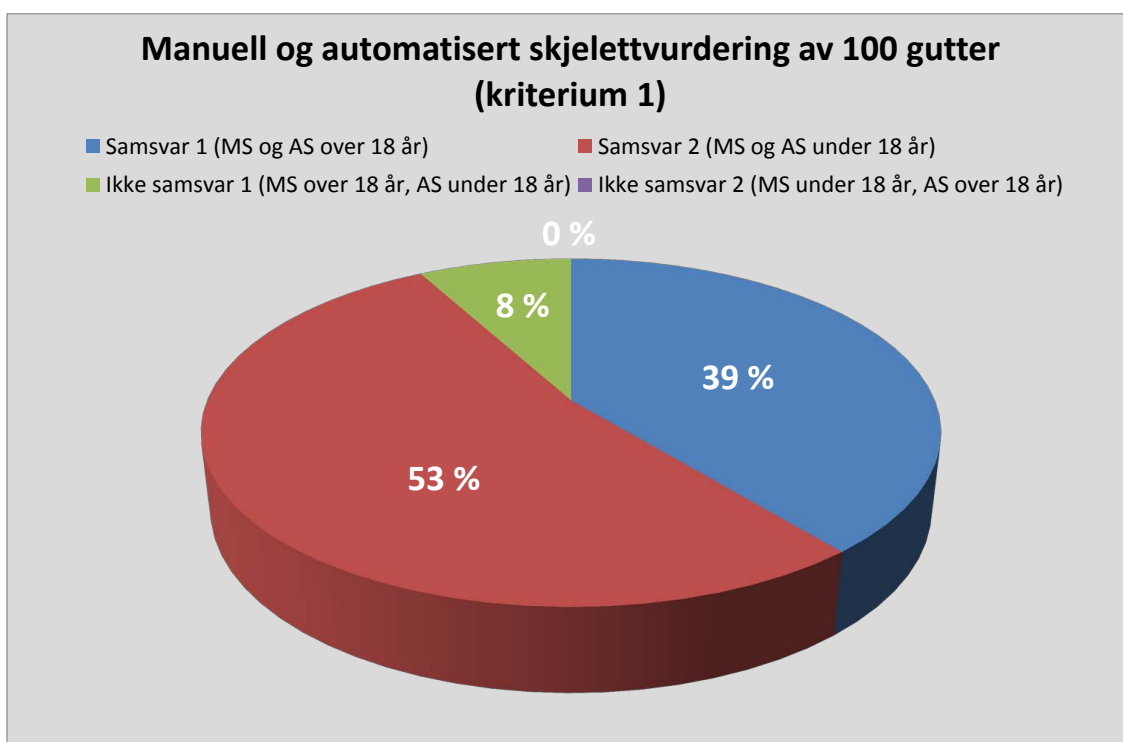


Figur 4. Manuell (øverst) og automatisert (nederst) skjelettvurdering utført på 100 utvalgte gutter. Manuell skjelettvurdering er gitt som heltall. Automatisert skjelettvurdering fra BoneXpert er gitt med to desimaler og derfor er gruppene gitt som halvåpne intervaller. Intervallet $[a,b)$ inneholder alle reelle tall som er større enn eller lik a og mindre enn b .

3.1 Kriterium 1

Vi lar kriteriet for at individet skal være 18 år eller eldre være likt for manuell og automatisert skjelettvurdering; skjelettalder fra de to aldersvurderingene må være 18 år eller høyere.

Kakediagrammet i figur 5 viser prosentandel gutter med og uten samsvar ifølge dette kriteriet. For 39% av observasjonene så er gutten 18 år eller eldre både fra manuell og automatisert skjelettvurdering (Samsvar 1) og for 53% av observasjonene så er gutten under 18 år både fra manuell og automatisert skjelettvurdering (Samsvar 2). Følgelig er det 92% samsvar mellom manuell og automatisert skjelettvurdering i henhold til om guttens alder vurderes til å være over eller under 18 år.



Figur 5. Prosentandel observasjoner med og uten samsvar fra aldersvurderinger av 100 utvalgte gutter. Samsvar er definert i henhold til om guttens alder vurderes til å være over eller under 18 år. Skjelettalder fra manuell skjelettvurdering (MS) og automatisert skjelettvurdering (AS) må være 18 år eller høyere for at individet skal være 18 år eller eldre.

Tabell 1 beskriver avvikene mellom manuell og automatisert skjelettvurdering for åtte gutter som ikke er i samsvar ifølge kriterium 1. Alle avvikene er positive, dvs. at guttene er 18 år eller eldre fra manuell skjelettvurdering og under 18 år fra automatisert skjelettvurdering (Ikke samsvar 1). Avvikene ligger mellom 0.12 år og 1.98 år. Gjennomsnittlig avvik for disse åtte guttene er 1.02 år og median avvik er 1.07 år².

2. Hans Henrik Thodberg har funnet ut at det var BoneXpert som gjorde feil skjelettvurdering av gutten med størst avvik (1.98 år). Gjennomsnittlig avvik for de syv andre guttene (avvikene ligger mellom 0.12 år og 1.27 år) er 0.88 år og median avvik er 1.05 år.

Tabell 1. Avvik mellom manuell og automatisert skjelettvrdering for kriterium 1. Åtte gutter er 18 år eller eldre fra manuell skjelettvrdering og under 18 år fra automatisert skjelettvrdering (Ikke samsvar 1).

Manuell	Automatisert	Avvik	Utfall
18	17.88	0.12	Ikke samsvar 1
18	17.55	0.45	
19	17.98	1.02	
19	17.95	1.05	
19	17.92	1.08	
19	17.84	1.16	
19	17.73	1.27	
19	17.02	1.98	

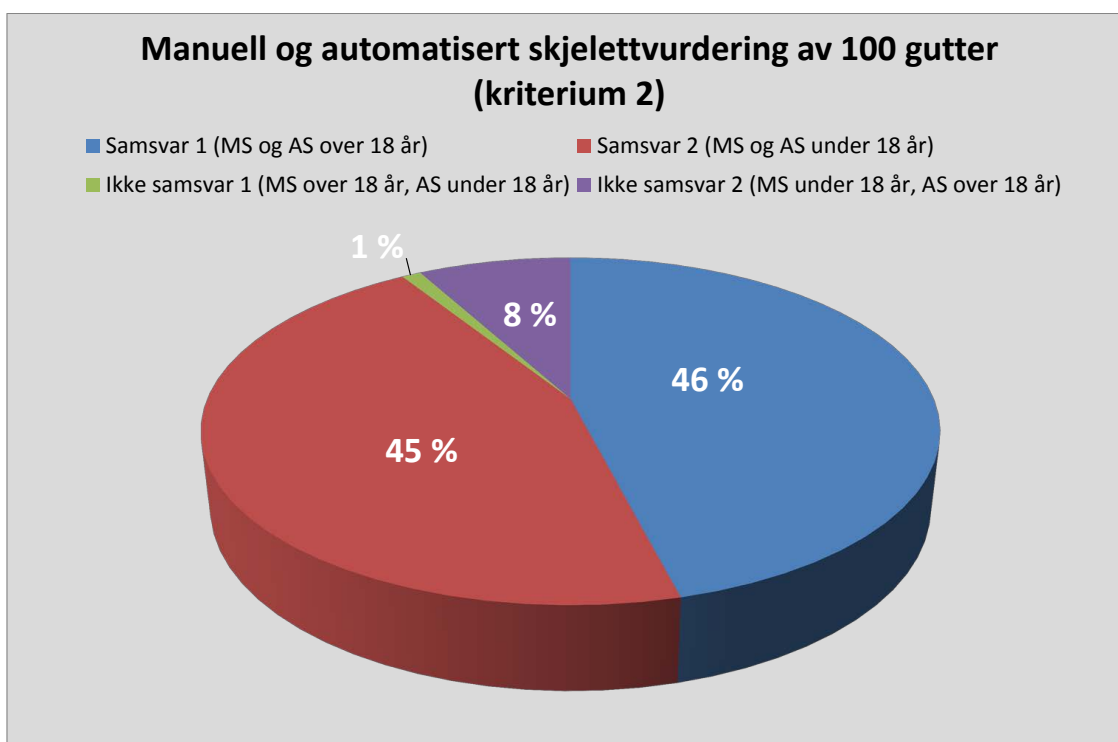
3.2 Kriterium 2

Vi ser på et annet kriterium for at individet skal være over eller under 18 år basert på den automatiserte skjelettvrderingen; skjelettalder fra BoneXpert må være 17.5 år eller høyere for at individet skal være 18 år eller eldre. Kriteriet for manuell skjelettvrdering er uendret, dvs. at skjelettalder fra manuell skjelettvrdering må være 18 år eller høyere for at individet skal være 18 år eller eldre.

Kakediagrammet i figur 6 viser prosentandel gutter med og uten samsvar ifølge dette kriteriet. For 46% av observasjonene så er gutten 18 år eller eldre både fra manuell og automatisert skjelettvrdering (Samsvar 1) og for 45% av observasjonene så er gutten under 18 år både fra manuell og automatisert skjelettvrdering (Samsvar 2). Følgelig er det 91% samsvar mellom manuell og automatisert skjelettvrdering i henhold til om guttens alder vurderes til å være over eller under 18 år.

Tabell 2 beskriver avvikene mellom manuell og automatisert skjelettvrdering for ni gutter som ikke er i samsvar ifølge kriterium 2. Det er åtte negative avvik der guttene er under 18 år fra manuell skjelettvrdering og 18 år eller eldre fra automatisert skjelettvrdering (Ikke samsvar 2). De negative avvikene ligger mellom -0.97 år og -0.54 år. Det er ett positivt avvik på 1.98 år der gutten er 18 år eller eldre fra manuell skjelettvrdering og under 18 år fra automatisert skjelettvrdering (Ikke samsvar 1). Gjennomsnittlig absoluttavvik for disse ni guttene er 0.81 år og median absoluttavvik er 0.60 år³.

3. Hans Henrik Thodberg har funnet ut at det var BoneXpert som gjorde feil skjelettvrdering av gutten med størst avvik (1.98 år). Gjennomsnittlig avvik for de åtte andre guttene (avvikene ligger mellom -0.97 år og -0.54 år) er -0.66 år og median avvik er -0.59 år.



Figur 6. Prosentandel observasjoner med og uten samsvar fra aldersvurderinger av 100 utvalgte gutter. Samsvar er definert i henhold til om guttens alder vurderes til å være over eller under 18 år. Skjelettalder fra manuell skjelettvrdering (MS) må være 18 år eller høyere og skjelettalder fra automatisert skjelettvrdering (AS) må være 17.5 år eller høyere for at individet skal være 18 år eller eldre.

Tabell 2. Avvik mellom manuell og automatisert skjelettvrdering for kriterium 2. Åtte gutter er under 18 år fra manuell skjelettvrdering og 18 år eller eldre fra automatisert skjelettvrdering (Ikke samsvar 2). Én gutt er 18 år eller eldre fra manuell skjelettvrdering og under 18 år fra automatisert skjelettvrdering (Ikke samsvar 1).

Manuell	Automatisert	Avvik	Utfall
17	17.97	-0.97	} Ikke samsvar 2
17	17.83	-0.83	
17	17.68	-0.68	
17	17.60	-0.60	
17	17.58	-0.58	
17	17.55	-0.55	
17	17.55	-0.55	
17	17.54	-0.54	
19	17.02	1.98	} Ikke samsvar 1

4 Oppsummering og konklusjon

Vi har utført en vellykket test av BoneXpert på 100 utvalgte røntgenbilder av gutter (EMA). Samlet vurdering er resultatet av den endelige medisinske aldersvurderingen, og er brukt til å se på aldersfordeling i utvalget. 95 av guttene har samlet vurdering på 15 år eller høyere og 67 av guttene har samlet vurdering på 17 år eller høyere.

Manuell skjelettvurdering har vært gitt som heltall og kriteriet for at individet skal være 18 år eller eldre har vært at skjelettalder fra manuell skjelettvurdering må være 18 år eller høyere. Automatisert skjelettvurdering fra BoneXpert har vært gitt med to desimaler. Kriteriene for at individet skal være 18 år eller eldre har vært at skjelettalder fra BoneXpert må være 18 år eller høyere (kriterium 1), eller 17.5 år eller høyere (kriterium 2).

Vi mener at det er godt samsvar mellom manuell og automatisert skjelettvurdering utført på Unilabs Norge, i henhold til om guttens alder vurderes til å være over eller under 18 år. For kriterium 1 er det 92% samsvar og for kriterium 2 er det 91% samsvar. Det gode samsvaret vi har sett i denne testen skyldes tre forhold, som alle er nødvendige for å få et godt resultat: (i) Den manuelle skjelettvurderingen er presist utført av Unilabs Norge, (ii) BoneXpert er presis og gir en pålitelig skjelettvurdering og (iii) kvaliteten på de analyserte røntgenbildene er god. Det er ingen undereksponering som gjør det vanskelig å lokalisere og tolke beinene i hånden i den automatiserte skjelettvurderingen.

Utvalget i vår test har vært relativt lite og det er behov for en større studie der BoneXpert anvendes på flere røntgenbilder for å verifisere det gode samsvaret. En større studie er også nødvendig for å kunne bestemme en grense for å avgjøre om individet er over eller under 18 år basert på den automatiserte skjelettvurderingen. Dersom kriteriet i vår test hadde vært at skjelettalder fra BoneXpert må være høyere enn 17.7 år for at individet skal være 18 år eller eldre, hadde vi fått 96% samsvar mellom de to skjelettvurderingene.

Uansett hvor grensen settes vil det, grunnet biologisk variasjon, være andeler feilklassifisering i begge retninger (Ikke samsvar 1 og Ikke samsvar 2). Med et større utvalg røntgenbilder av individer med kjent kronologisk alder kan disse feilratene beregnes for ulike grenser. Ved å veie ulike hensyn kvantitativt opp mot hverandre kan feilratene i begge retninger balanseres til akseptable nivåer og det kan bestemmes en operativ standard for å avgjøre om individet er over eller under 18 år.

Anerkjennelse

Vi ønsker å takke administrerende direktør i Visiana, Hans Henrik Thodberg, for bistand med installering og bruk av BoneXpert. Thodberg har også bidratt med beskrivelser av programvaren som vi har testet, en prototype av BoneXpert med versjon nummer 2.1.1.1, og diskusjoner i forbindelse med å fastsette et kriterium for å avgjøre om individet er over eller under 18 år.

Referanser

- Gelbrich, B., Frerking, C., Weiß, S., Schwerdt, S., Stellzig-Eisenhauer, A., Tausche, E. og Gelbrich, G. (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. *Annals of Human Biology*, 42(4):389–396.
- Greulich, W. W. og Pyle, S. I. (1959). *Radiographic Atlas of Skeletal Development of the Hand and Wrist*. Stanford, CA: Stanford University Press, 2nd edition.
- Haugen, M., Grøgaard, J., Kjellevand, T. O. og Kvaal, S. I. (2016). Sammenligning av to metoder for aldersvurdering av enslige mindreårige asylsøkere. NR-Note, SAM-BA/05/16.
- Martin, D. D., Heil, K., Heckmann, C., Zierl, A., Schaefer, J., Ranke, M. B. og Binder, G. (2013). Validation of automatic bone age determination in children with congenital adrenal hyperplasia. *Pediatric Radiology*, 43(12):1615–1621.
- Thodberg, H. H. (2009). Clinical review: An automated method for determination of bone age. *J. Clin. Endocrinol. Metab*, 94(7):2239–2244.
- Thodberg, H. H., Kreiborg, S., Juul, A. og Pedersen, K. D. (2009). The BoneXpert method for automated determination of skeletal maturity. *IEEE Trans Medical Imaging*, 28(1):52–66.
- Thodberg, H. H. og Sävendahl, L. (2010). Validation and reference values of automated bone age determination for four ethnicities. *Academic Radiology*, 17(11):1425–1432.
- van Rijn, R. R., Lequin, M. H. og Thodberg, H. H. (2009). Automatic determination of Greulich and Pyle bone age in healthy Dutch children. *Pediatric Radiology*, 39(6):591–597.