

REPORT

2018

A SYSTEMATIC REVIEW:

Development stages of the knee and ankle by computed tomography and magnetic resonance imaging for estimation of chronological age

Title Development stages of the knee and ankle by computed tomography and magnetic resonance imaging for estimation of chronological age: a systematic review

Norwegian title Utviklingsstadier av kne og ankel med computertomografi og magnetresonanstomografi for å estimere kronologisk alder: en systematisk oversikt

Publisher Norwegian Institute of Public Health
(Folkehelseinstituttet)
Camilla Stoltenberg, Director General

Authors Kristoffer Yunpeng Ding, project leader, *Norwegian Institute of Public Health*
Pål Skage Dahlberg, *Oslo University Hospital*
Veslemøy Rolseth, *Oslo University Hospital*
Annhild Mosdøl, *Norwegian Institute of Public Health*
Gyri Hval Straumann, *Norwegian Institute of Public Health*
Øyvind Bleka, *Oslo University Hospital*
Gunn Elisabeth Vist, *Norwegian Institute of Public Health*

ISBN 978-82-8082-947-4

Type of report Systematic review

No. of pages 44 (58 including appendices)

Client Norwegian Institute of Public Health

Subject heading (MeSH) Age assessment, knee and ankle

Citation Ding KY, Dahlberg PS, Rolseth V, Mosdøl A, Straumann GH, Bleka Ø, Vist GE. Development stages of the knee and ankle by computed tomography and magnetic resonance imaging for estimation of chronological age: a systematic review (Utviklingsstadier av kne og ankel med computertomografi og magnetresonanstomografi for å estimere kronologisk alder: en systematisk oversikt). Report 2018. Oslo: Norwegian Institute of Public Health, 2018.

Table of contents

Table of contents	2
Key messages	4
Executive summary	5
Hovedbudskap (norsk)	7
Sammendrag (norsk)	8
Preface	10
Introduction	11
Method	17
Inclusion criteria.....	17
Literature search.....	18
Article selection and assessment.....	18
Risk of bias and data extraction	18
Statistical analyses.....	19
GRADE framework.....	19
Results	20
Results of literature search.....	20
Excluded studies.....	21
Separate list.....	21
Description of included studies.....	21
Risk of bias assessment of included studies according to QUADAS-2.....	22
Population distribution in included studies.....	23
Presentation of ossification stages.....	25
Chronological age distribution and distal femoral ossification stages	25
Chronological age distribution and proximal tibial ossification stages.....	31
Chronological age distribution and ankle ossification stages.....	36
Kappa agreement coefficients.....	37

Discussion 38

 Summary and key findings..... 38

 Quality of the results..... 38

 Future perspective..... 40

 Strengths and limitations..... 40

Conclusion 42

References 43

Appendices 45

 Appendix 1. Literature search strategy..... 45

 Appendix 2. Description of included studies with quality assessment..... 49

 Appendix 3. Separate list..... 53

 Appendix 4. Excluded studies 54

 Appendix 5. Abbreviations 56

 Appendix 6. Project protocol..... 57

Key messages

Forensic age estimation of adolescents is an important field of research and practice to ensure that unaccompanied, young asylum seekers receive their rights and adults are not treated as minors. In this systematic review, we summarized evidence regarding age estimation on knee and ankle ossification using computed tomography (CT) and magnetic resonance imaging (MRI).

We found no relevant studies using CT, but four using MRI. These studies were from France, Germany and Turkey, and included 1250 participants between 8 to 30 years old. Two different classification methods were presented for knee ossification, and one method for ankle ossification. All three methods showed good intra- and inter-observer reliability. However, most of the studies were conducted with limited number of participants in each ossification stage and/or had uneven number of participants in each age group, which may lead to substantial variation in chronological age distribution in each development stage.

Given the limited and potentially biased results, we decided not to conduct meta-analysis. More studies with sufficient sample size and a uniform age structure are warranted for more accurate and reliable age estimation using these methods.

Title: Development stages of the knee and ankle by computed tomography and magnetic resonance imaging for estimation of chronological age: a systematic review.

Type of publication:
Systematic review

A review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies.

Doesn't answer everything:

- Excluded studies are not evaluated
- No recommendation
- No cost-effectiveness

Publisher:
Norwegian Institute of Public Health

Updated:
Last search for studies: April 2017

Peer review:
- Kjetil G. Brurberg (internal)
- Lene K. Juvet (internal)
- Øyvind Melien (internal)
- Sue Black (University of Dundee)
- Lucina Hackman (University of Dundee)

Executive summary

Introduction

Age estimation of living adolescents and young adults has become increasingly important in modern society, especially for asylum seekers who come to Norway without legal documentation of their chronological age. It is therefore necessary to assess chronological age in forensic practice to ensure that children receive their entitled rights and that adults are not treated as minors. Age assessments based on skeletal development of left hand-wrist and third molar teeth using radiographs have been used in Norway for years. We have previously published systematic reviews assessing the agreement between chronological age and skeletal development with the Greulich & Pyle atlas for left hand-wrist radiographs, the Demirjian's stages for the third molar teeth, and computed tomography (CT) and magnetic resonance imaging (MRI) for medial clavicle ossification, respectively. However, methods using knee and ankle ossification for age estimation have not been reviewed. We therefore conducted a systematic review to evaluate the evidence of chronological age distribution using identified methods of knee and ankle ossification by CT and MRI.

Method

We searched for studies in the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, Embase and Google Scholar. This was a joint search conducted for studies using radiographs of left hand-wrist, third molar teeth, CT and MRI of the medial clavicle, knee and ankle in both males and females. An update literature search was conducted in April 2017 for clavicle, knee and ankle. Two of the authors screened the literature independently from the title and abstract first, and subsequently full-text screening. We included studies that presented age distribution according to knee and ankle ossification stages. Two of the authors independently assessed the risk of bias and applicability of the included studies based on the QUADAS-2 checklist. The findings were summarized as forest plots. The mean chronological age and the standard deviation were summarized in each ossification stage for males and females, respectively.

Results

We found 10059 references in the first literature search and 663 in the second. In total 28 potentially relevant publications were forwarded for full-text screening. We did not find any CT studies matching the inclusion criteria. Four MRI studies were included. The included studies were from France (one study), Germany (one study) and Turkey (two studies). These studies were published from 2012 to 2016, involving 1250 participants from 8 to 30 years old. Two different methods were reported for knee ossification (Dedouit's method and Kramer's method) and one for ankle ossification (Saint-Martin's

method). Most of the studies showed good intra- and inter-observer reliability ($K > 0.80$). However, all of the included studies showed high risk of age mimicry bias due to uneven number of participants in each age group. Besides, three of the included studies had relatively small sample size in most of the age groups ($n < 15$), leading to large variation in age estimation. Most of the pooled estimates of chronological age and the distribution in knee and ankle ossification showed high heterogeneity ($I^2 > 75\%$).

Discussion

We only found four studies meeting our inclusion criteria, yielding limited evidence on age estimation by knee and ankle ossification methods. Besides, the results from the included studies suffered from relatively small sample size and uneven number of participants in each age group, leading to substantial variation between studies. Nevertheless, these findings are in line with our previous systematic reviews on age estimation based on the development of third molar teeth and medial clavicle, revealing the potentially high risk of age mimicry bias and uncertainty associated with the assessments in age estimation studies.

Conclusion

Three MRI classification methods were used in four studies describing chronological age distribution in knee and ankle ossification, which yields very limited evidence for us to assess the validity of these methods. Furthermore, the majority of the included studies had limited and uneven sample size in each age group, leading to substantial variation on observed age distribution across studies. We did not find any relevant studies using CT classification method on knee or ankle ossification. Future studies with even numbers in each age group, wide age spectrum and sufficient sample size are warranted for a better understanding of the age distribution in the knee and ankle ossification stages.

Hovedbudskap (norsk)

For å sikre at enslige, unge asylsøkere får de rettigheter de har krav på og at voksne ikke behandles som mindreårige, har aldersestimering av ungdom blitt et viktig felt innen både forskning og praksis. I denne systematiske oversikten oppsummerer vi den forskningsbaserte dokumentasjonen om aldersestimering basert på forbeningen i kne og ankel målt med computertomografi (CT) eller magnetresonanstomografi (MR).

Vi fant ingen relevante studier som brukte CT, men fire basert på MR. Disse studiene var fra Frankrike, Tyskland og Tyrkia, og omfattet totalt 1250 deltakere i alderen fra 8 til 30 år. To ulike klassifiseringsmetoder ble brukt for forbeningen i kne, og én metode for forbeningen i ankel. Alle de tre metodene hadde godt intra- og inter-observatør samsvar. Det fleste studiene hadde imidlertid begrenset antall deltakere i hvert forbeningsstadium eller hadde et ujevnt antall deltakere i hver aldersgruppe, som kan føre til betydelig variasjon av observert kronologisk alder i hvert forbeningsstadium.

Gitt de begrensede og potensielt skjevframstilte resultatene, utførte vi ikke metaanalyser. Flere studier med tilstrekkelig antall deltakere og en jevn aldersfordeling er nødvendig for mer eksakt og pålitelig aldersestimering basert på disse metodene.

Titel: Utviklingsstadier av kne og ankel med computertomografi og magnetresonanstomografi for å estimere kronologisk alder: en systematisk oversikt.

Publikasjonstype:

Systematisk Oversikt

En systematisk oversikt er resultatet av å

- innhente
- kritisk vurdere og
- sammenfatte relevante forskningsresultater ved hjelp av forhåndsdefinerte og eksplisitte metoder.

Svarer ikke på alt:

- Ingen studier utenfor de eksplisitte
- inklusjonskriteriene
- Ingen helseøkonomisk evaluering
- Ingen anbefalinger

Hvem står bak denne publikasjonen?

Folkehelseinstituttet

Når ble litteratursøket utført?

Siste søket: April 2017

Fagfeller:

- Kjetil G. Brurberg (intern)
- Lene K. Juvet (intern)
- Øyvind Melien (intern)
- Sue Black (Universitetet i Dundee, UK)
- Lucina Hackman (Universitetet i Dundee, UK)

Sammendrag (norsk)

Innledning

Aldersestimering av ungdom og unge voksne har i økende grad blitt viktig i moderne samfunn. Hvert år kommer unge asylsøkere til Norge uten juridisk dokumentasjon på fødselsdato. Det er nødvendig å fastsette deres kronologiske alder for å sikre at barn får de rettighetene de har krav på, og for at voksne ikke behandles som mindreårige. I Norge har alder på asylsøkere blitt vurdert ved å evaluere modningen av hånd-håndrot skjelett og tannutviklingen. Vi har tidligere publisert systematiske oversikter som vurderte samsvar mellom kronologisk alder og henholdsvis (I) skjelettalder av hender med Greulich & Pyle atlaset, (II) Demirjians metode for å vurdere visdomstennene (tredje molar), og (III) computertomografi (CT) og magnetresonanstomografi (MR) for forbening i kragebein. Metoder basert på forbeningen i kne og ankel har imidlertid ikke vært oppsummert. Vi utførte en systematisk oversikt for å vurdere forskningsdokumentasjon om fordeling av kronologisk alder fra utviklingsstadier av kne og ankel basert på CT og MR.

Metode

Vi søkte etter studier i Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, Embase og Google Scholar. Et felles søk ble gjort for å finne studier som benyttet røntgen av tenner og hånd, CT og MR av det mediale kragebein, kne og ankel, hos individer av begge kjønn mellom 10 og 35 år. Et oppdatert litteratursøk ble gjennomført i april 2017 kun på CT og MR av kragebein, kne og ankel. To av forfatterne vurderte uavhengig av hverandre referansene basert på tittel og sammendrag, og siden som fulltekst screening. Vi inkluderte studier som presenterte aldersfordeling for forbening i kne og ankel. To av forfatterne vurderte uavhengig av hverandre risiko for systematisk skjevheter basert på QUADAS-2 sjekklisten. Funnene ble presentert i forestplott. Gjennomsnittlig kronologisk alder og standardavvik ble oppsummert for hvert forbeningstadium, separat for gutter og jenter.

Resultat

Vi fant 10059 referanser i det første litteratursøket og 663 i det andre. Totalt 28 potensielt relevante studier ble vurdert i fulltekst. Vi fant ingen CT studier som oppfylte inklusjonskriteriene. Fire MR studier ble inkludert. Studiene var utført i Frankrike (én studie), Tyskland (én studie) og Tyrkia (to studier). Disse studiene var publisert fra 2012 til 2016, og involverte 1250 deltakere i alderen fra 8 til 30 år. To ulike klassifiseringsmetoder ble presentert for forbening i kne for aldersestimering (Dedouit's metode og Kramer's metode) og en metode for ankel (Saint-Martin's metode). De fleste metodene viste godt intra- og inter-observatør samsvar ($K > 0.80$). Imidlertid hadde alle de inkluderte studiene høy risiko for aldersmimikering på grunn av ujevnt antall deltakere i hver aldersgruppe. Tre studier hadde relativt lavt

antall deltakere i hver aldersgruppe ($n < 15$), som gir betydelig variasjon i observert kronologisk alder. Følgelig ansees ikke samlede estimat for kronologisk alder som pålitelige, og fordelingen av kronologisk alder for hvert stadium viste stor heterogenitet ($I^2 > 75\%$).

Diskusjon

Vi fant svært begrenset tilgjengelig informasjon om aldersestimering basert på forbening i kne og ankel. Resultatene fra de inkluderte studiene var begrenset av relativt lavt antall deltakere og ujevnt antall deltakere i hver aldersgruppe, som bidrar til stor variasjon mellom studiene. Disse resultatene er i samsvar med funnene fra våre tidligere systematiske oversikter om estimering av kronologisk alder ved hjelp av utvikling av visdomstenner og forbening av det mediale kragebeinet, som også viste risiko for aldersmimikering.

Konklusjon

De tre MR studiene gir svært begrenset informasjon for å vurdere validiteten av disse metodene. Majoriteten av de inkluderte studiene hadde et begrenset antall deltakere eller ulikt antall i hver aldersgruppe som gir betydelig variasjon i observert fordeling av alder. Framtidige studier med likt antall deltakere i hver aldersgruppe, et vidt aldersspektrum og et stort utvalg, er nødvendig for å gi en bedre forståelse av aldersfordeling i forbeningsstadier i kne og ankel. Det mangler gode studier med bruk av CT og MR for å vise aldersfordeling for utviklingsstadier av kne og ankel.

Preface

This systematic review summarizes evidence of age distribution according to the ossification stages of knee and ankle by computed tomography and magnetic resonance imaging. We have previously published two systematic reviews on age assessments of left hand-wrist radiographs using Greulich & Pyle atlas (1) and third molar teeth using Demirjian's stages (2), respectively. Another systematic review on age assessment of medial clavicle epiphysis is forthcoming. We have chosen to write all four reports as separate documents but use consistent texts throughout the documents where relevant.

The project group consisted of:

- Kristoffer Y. Ding, project leader, Norwegian Institute of Public Health
- Annhild Mosdøl, Norwegian Institute of Public Health
- Pål Skage Dahlberg, Oslo University Hospital
- Veslemøy Rolseth, Oslo University Hospital
- Gyri H. Staumann, Norwegian Institute of Public Health
- Øyvind Bleka, Oslo University Hospital
- Gunn E. Vist, Norwegian Institute of Public Health

We thank Kjetil Gundro Brurberg, Lene Kristine Juvet Øyvind Melien for being the internal reviewers, Professor Sue Black (University of Dundee) and Doctor Lucina Hackman (University of Dundee) as the external reviewers for our research protocol and final report. We also want to thank Marit Johansen as peer reviewer for our literature search.

Notably, Figure 1 is reprinted from “Age assessment by magnetic resonance imaging of the knee: a preliminary study”, Epub 2011 Dec 5, Dedouit F. et al, Copyright (2012), with permission from Elsevier [Forensic Science International].

Figure 2 is reprinted by permission from Springer: Int J Legal Med [Kramer, J.A., et al., Forensic age estimation in living individuals using 3.0 T MRI of the distal femur. Int J Legal Med, 2014. 128(3): p. 509-14, [COPYRIGHT] (2014).

Figure 3 is reprinted by permission from Springer: Int J Legal Med [Saint-Martin, P., et al., Age estimation by magnetic resonance imaging of the distal tibial epiphysis and the calcaneum. Int J Legal Med, 2013. 127(5): p. 1023-30, [COPYRIGHT] (2013).

Kåre B. Hagen
Director Reviews and HTA

Gunn E. Vist
Unit leader

Kristoffer Y. Ding
Project leader

Introduction

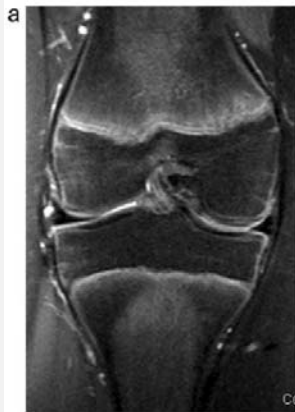
Age estimation of living adolescents and young adults with unknown age has been of considerable interest in forensic practice and research, especially for purposes of age assessment of refugees (3). There are a number of observable biological changes as a person grows and the body develops. The assessment of specific developmental stages constitutes the basis for medical age estimation. Widely used approaches for age estimation of living adolescents and young adults are the development of left hand-wrist radiographs and third molar teeth (1, 2). Besides, analysis of the medial clavicle epiphysis may also be recommended, especially for the assessment of age 18 and older (4-6).

From 1 January 2016, the Department of Forensic Sciences at the Norwegian Institute of Public Health, now at the Oslo University Hospital, received the assignment to take national responsibility for medical age assessment. It was decided to conduct systematic reviews of various methods used for medical age assessments. Thus far, we have summarized results of age estimation of the left hand-wrist radiographs using Greulich-Pyle atlas (1) and third molar teeth using Demirjian's stages (2). A systematic review focusing on age assessment of the medial clavicular epiphysis is forthcoming. These systematic reviews revealed considerable heterogeneity between included studies on the agreement of age distribution according to maturation stages. The heterogeneity has often been explained by regional and phenotype differences between population groups (7, 8). However, an additional explanation may come from a particular selection bias, called "age mimicry", which has been discussed in detail earlier (1, 2). In short, age mimicry occur when the sizes of each age group of the enrolled population are uneven, or having insufficient age range to cover the studied development stages. These factors affect the results, which will "mimic" the age distribution of the included population by over- or under-representation of certain ages.

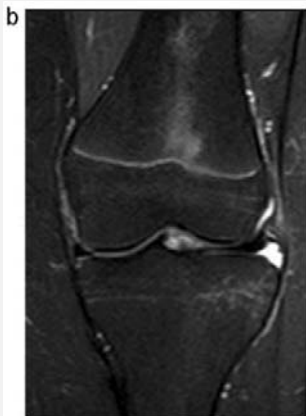
The purpose of this systematic review is to summarize the results of age estimation with knee and ankle ossification stages by means of computed tomography (CT) and magnetic resonance imaging (MRI), since CT and MRI are often applied to assess knee and ankle ossification. The current systematic review excluded studies using radiographs because two-dimensional radiographs from X-rays do not always depict the ossification of the knee in a clear and consistent manner (9). The current review, together with the other systematic reviews, constitutes part of a basis for further discussion and recommendation of how medical age assessments should be conducted in Norway.

Description of ossification stages of knee

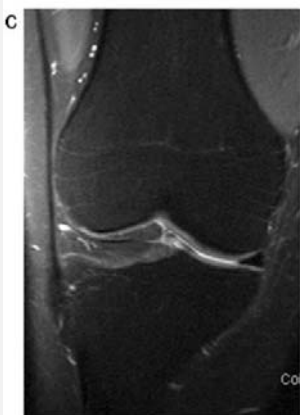
Two methods were reported to evaluate the ossification stages of knee by MRI. Dedouit et al. (10) developed an original MRI staging system for epiphyseal fusion of growth plate maturation of the knee (Figure 1). Both distal femoral and proximal tibial epiphyses were separately evaluated in these stages. Five growth plate patterns were designated as stages I–V:



Stage I (a): continuous horizontal cartilage signal intensity present between the metaphysis and the epiphysis, stripe-like, with a thickness greater than 1.5 mm and a multilaminar appearance. The multilaminar appearance was seen as decreased signal intensity in the upper layer, increased signal intensity in the middle layer, and decreased signal intensity in the lower layer.



Stage II (b): continuous horizontal linear cartilage signal intensity present between the metaphysis and the epiphysis, with a thickness greater than 1.5 mm, with increased signal intensity but without a multilaminar appearance.

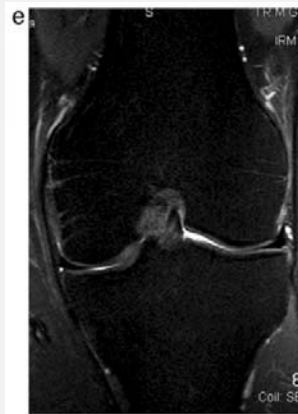


Stage III (c): continuous horizontal linear cartilage signal intensity present between the metaphysis and the epiphysis, with a thickness less than 1.5 mm, with increased signal intensity.

Figure 1. Coronal MR scans of the knee (Dedouit's method).



Stage IV (d): discontinuous horizontal linear cartilage signal intensity present between the metaphysis and the epiphysis, with a thickness less than 1.5 mm, with discontinuous increased signal intensity.



Stage V (e): no increased signal intensity between the metaphysis and the epiphysis.

Figure 1 (continued). Coronal MR scans of the knee (Dedouit's method).

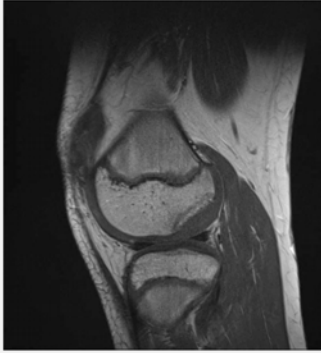
The second method, by Kramer et al. (11), modified a classification method on medial clavicle epiphysis that was proposed by Schmeling et al (5) and by Kellinghaus et al. (12) whereby the ossification stages of distal femur were defined as follows (no MR images were shown for stage 1, 2a, 2b and 5):

Stage 1: The epiphysis has not yet ossified.

Stage 2a: The length of the ossified epiphysis is one third or less compared to the width of the metaphyseal ending.

Stage 2b: The length of the ossified epiphysis is between one third and two thirds compared to the width of the metaphyseal ending.

Figure 2. MRI of distal femur epiphysis (Kramer's method)



Stage 2c: The length of the ossified epiphysis is over two thirds compared to the width of the metaphyseal ending.



Stage 3a: Epiphyseal-metaphyseal fusion completes one third or less of the former gap between epiphysis and metaphysis.

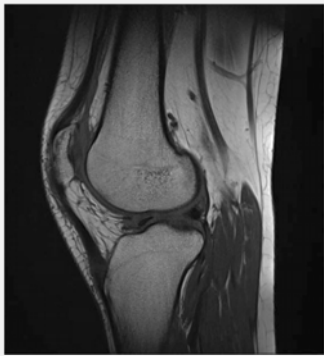


Stage 3b: Epiphyseal-metaphyseal fusion completes between one third and two thirds of the former gap between epiphysis and metaphysis.



Stage 3c: Epiphyseal-metaphyseal fusion completes over two thirds of the former gap between epiphysis and metaphysis.

Figure 2 (continued). MRI of distal femur epiphysis (Kramer's method)



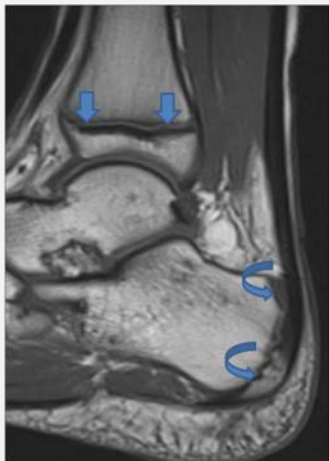
Stage 4: The epiphyseal cartilage is fully ossified, and the epiphyseal scar is visible.

Stage 5: The epiphyseal cartilage has fused completely, and the epiphyseal scar is no longer visible.

Figure 2 (continued). MRI of distal femur epiphysis (Kramer's method)

Description of ossification stages of ankle

The method of ossification stages of ankle, introduced by Ekizoglu et al. (13), used three-stage scoring system that was developed by Saint-Martin et al. (14) as follows:



Stage 1: No fusion. A gap is present between the metaphysis and epiphysis with total absence of bridging (indicated by blue arrows). On the superior part of the gap, there is a continuous horizontal (sometimes multilaminar) stripe with cartilage-like signal intensity.

Figure 3. The metaphysis and epiphysis of the distal tibia and calcaneus



Stage 2: *Partial fusion. The gap between the metaphysis and epiphysis is not continuous (indicated by blue arrows). In at least one image of the series, there is a hazy area instead of the horizontal stripe, showing bridging between the epiphysis and metaphysis. Both parts are still totally separated in some portions and in at least one image of the series.*



Stage 3: *Complete fusion. The epiphysis and metaphysis are united in all images of the series. A thin horizontal line may remain in some cases, representing the epiphyseal scar.*

Figure 3 (continued). The metaphysis and epiphysis of the distal tibia and calcaneus

Method

The current project included systematic literature searches on studies focusing on age estimation of knee and ankle by CT and MRI. This systematic review is based on our methodological guideline “*Slik oppsummerer vi forskning*” published by the former Norwegian Knowledge Center (15). We used the following specifications:

Inclusion criteria

Study design:	Studies that summarized data on chronological age distribution of knee and ankle ossification using CT and MRI
Population:	Living adolescents and young adults between 10 and 25 years old with no pathological problems on knee and ankle
Index test:	Ossification stages of knee (distal femur and proximal tibia) and ankle (distal tibia and calcaneus) by CT and MRI
Reference:	Confirmed chronological age
Outcome:	Chronological age summary (mean \pm SD) in each stage
Language:	Language limits was not applied to the searches. However, project members only read Chinese, Danish, English, German, Japanese, Norwegian, and Swedish. Relevant publications that could not be read due to language limitation would have been listed in a separate table

Exclusion Criteria:

- Studies without full-text (conference abstracts)
- Studies that is not an empirical study
- Studies that have included dead individuals instead of living human beings
- Studies that focused on osteometric parameters instead of age estimation
- Studies with less than 50 participants between 10 and 25 years old

Separate list

After reading full-text articles, we built a separate list for future reference to track studies that might have relevant data, but did not present the data in a way that we could utilize for this systematic review.

Literature search

Research librarian Gyri Hval Straumann created and conducted the literature searches, librarian Marit Johansen peer-reviewed the search strategies. We searched for primary studies with no limit of study design, publication time or language in the following databases:

- Cochrane Central Register of Controlled Trials (CENTRAL)
- MEDLINE (Ovid) and Pubmed [sb]
- Embase (Ovid)
- Google Scholar

The first search was carried out on 19 May 2016. This was a joint search conducted for studies using radiographs of teeth and hand-wrist, CT and MRI of the medial clavicle, knee and ankle in both males and females. An update search was conducted in April 2017 only for clavicle, knee and ankle. The search strategies are presented in Appendix 1.

Article selection and assessment

For the first search, six review authors (AH, GEV, GHS, KYD, PSD and VR) independently screened abstracts identified by the searches; for the second search, two review authors screened abstracts independently (GEV and KYD). All abstracts were screened in duplicates via the systematic reviews web-application Rayyan (16).

Articles were excluded if the title and/or abstract did not meet the inclusion criteria. For potentially relevant studies, the full-text articles were obtained and screened by two reviewers independently, with discrepancies resolved by consensus of reviewers. Studies that were considered as relevant to the review topic but did not meet all the inclusion criteria for the review were listed in the 'Characteristics of excluded studies' table, with the reason for their exclusion described. We recorded the selection process in sufficient detail to complete a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

Risk of bias and data extraction

To evaluate the risk of bias (methodical quality) of included studies, we adopted a revised QUADAS-2 checklist that has been described in detail in the previous age estimation projects on hand-wrist and teeth (1, 2). Notably, this checklist takes particular concern on age mimicry.

KYD extracted the following information from articles, and GEV double-checked the data accordingly:

- Where and when the study carried out (country and year)
- Scoring method

- Age range, sex, and sample size (for evaluating age mimicry bias)
- Study design
- Sample selection method

In addition, we extracted the following data for data synthesis:

- Mean age and its standard deviation (SD) of chronological age for each stage
- Total number of participants in each stage
- Kappa agreement coefficient

Statistical analyses

Our primary outcome is the mean age and its 95% confidence interval (CI) distributed in each ossification stage and substage of knee and ankle. We illustrated the results with forest plot, which would also include the information about heterogeneity. A secondary outcome is the kappa agreement coefficient in given stages to illustrate the intra- and inter-observer variability for different methods.

Heterogeneity between included studies was evaluated by Cochran's Q test, P value of 0.10 is used to determine statistical significance. In addition, I^2 is calculated for quantifying inconsistency of results between studies. Summarized results with high heterogeneity need to be interpreted with caution. The analyses was performed using R program (version: 3.4.0) with package "metafor" (17).

GRADE framework

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) tool is often used in systematic reviews to rate the quality and certainty of the included evidence. However, the current systematic review is not a typical diagnostic accuracy assessment study, where one presents positive/negative results with sensitivity and specificity analysis. Therefore, evaluating evidence quality by GRADE could not be conducted in the current systematic review.

Results

Results of literature search

We initially searched electronic database and registry in May 2016 (search I), and found 10059 potential relevant publications on age estimation using hand-wrist, teeth, clavicles, knees and ankles after removing duplicates. Among those publications, we found 25 potential relevant articles on age estimation utilizing knee and ankle MRI. In addition, we carried out an update search in April 2017 and found 663 articles (search II), in which three articles were of interest and were read in full-text. We did not find any study using CT. Four MRI studies were included (three on knee and one on ankle). Process in detail is described below (Fig. 4).

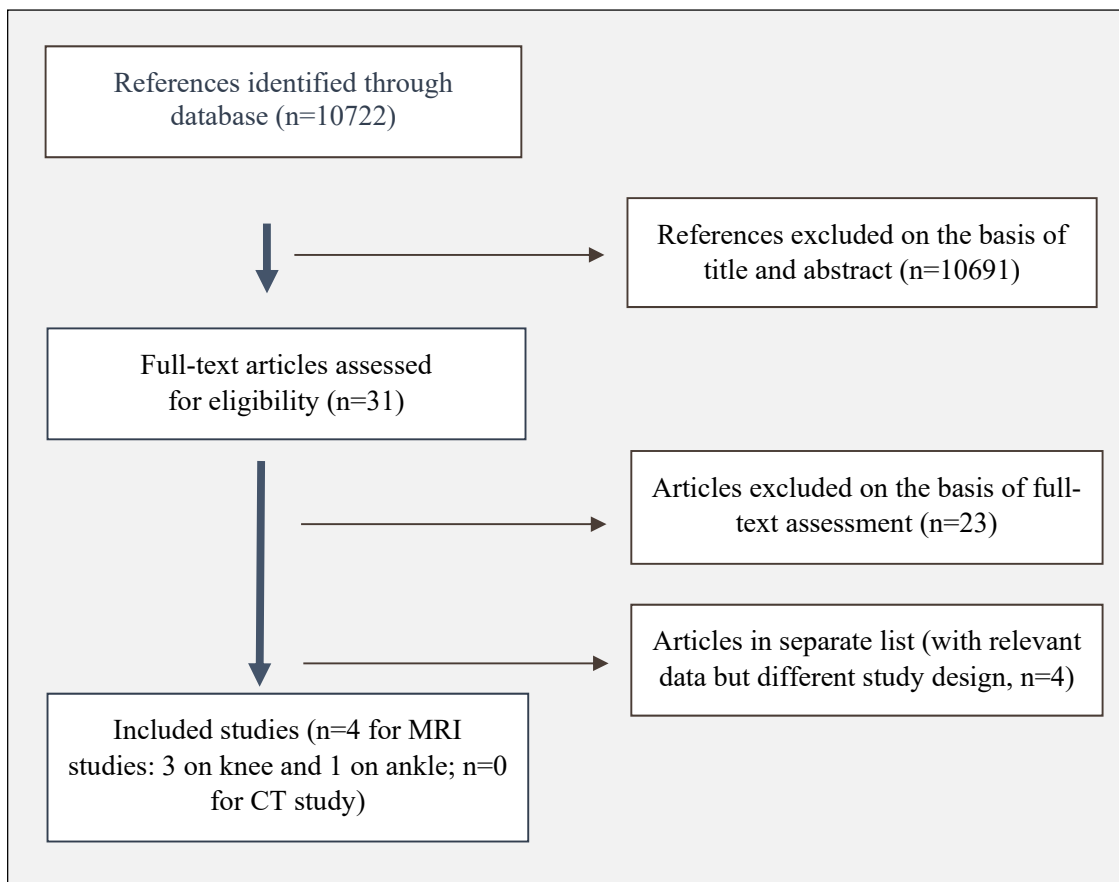


Figure 4. Flow chart of literature selection

Excluded studies

Of the 31 studies obtained in full-text, we excluded 23 studies due to conflicts against the inclusion criteria. See Appendix 4 for a list of excluded studies with reasons for exclusion.

Separate list

We put four MRI studies on a separate list. The purpose of this list is to track potential relevant studies for future reference. These studies appeared to have data that we might be able to use, but the design or the format of the results could not be incorporated into the current analyses. Among the four studies, two assessed the chronological age distribution by MRI of the knee, and two articles assessed the chronological age distribution by MRI of the ankle. The characteristics of the studies on this separate list can be found in Appendix 3.

Description of included studies

The four MRI studies that we included in the systematic review have adopted two methods on evaluating the ossification stages of the knee and one method on the ankle. The characteristics of the included studies are summarized in Table 1.

Table 1. Characteristics of included studies

Author	Year	Country	Sex	Stages	Size	Age range (min-max)
Knee						
Dedouit	2012	France	Both	1-5	290	10-30
Kramer	2014	Germany	Both	1-5	290	10-30
Ekizoglu	2016	Turkey	Both	1-5	503	10-30
Ankle						
Ekizoglu	2015	Turkey	Both	1-3	167	8-25

Studies included in the project were published from 2012 to 2016. The three studies on knee ossification are from France, Germany and Turkey. We only found one study that has assessed the ossification stages of ankle with MRI, which was published in 2016 from Turkey. The MRI settings and relevant parameters are summarized in Table 2.

Table 2. MRI settings of included studies

Study	Method	Sequence	Tesla
Knee			
Dedouit 2012	Dedouit	T2	1.5T
Ekizoglu 2016	Dedouit	T2	3T
Kramer 2014	Kramer	T1	3T
Ankle			
Ekizoglu 2015	Saint-Martin	T1	1.5T

The complete sample size for the knee assessment was 1083, which covered both genders from 10-30 years. The ankle assessment was comprised of 167 participants from both genders between 8 to 25 years.

Risk of bias assessment of included studies according to QUADAS-2

We summarized the information on the quality assessment of the included studies based on the QUADAS-2 checklist (18).

Table 3. Quality evaluation of the included studies based on QUADAS-2 checklist, with additional consideration of age mimicry.

Author, year	Domains for quality evaluation based on QUADAS-2				
	Selection bias		Index test interpretation	Reference standard	Patient flow and timing bias
	Patient selection	Age mimicry			
Dedouit, 2012					
Kramer, 2014					
Ekizoglu, 2016					
Ekizoglu, 2015					

Low risk, unclear risk and high risk for systematic bias in the included study.

All the included studies showed low risk of bias in Patient selection, Reference standard, and Patient flow and timing bias. Two studies (Kramer 2014 and Ekizoglu 2015) showed unclear risk on Index test interpretation. In addition, with age range and corresponding size in each age group, we could assess whether the distribution of chronological age for given developmental stages were prone to the influence of age mimicry bias. All the included studies showed high risk of age mimicry bias due to unevenly distributed number of subjects in some age groups. Notably, three of the studies (Dedouit 2012, Kramer 2014 and Ekizoglu 2015) included limited number of participants in most of the age groups ($n < 10$). The detailed information on the assessment of risk of bias using QUADAS-2 can be found in Appendix 2.

Population distribution in included studies

All four studies reported their sample size in each age group for both genders. We have summarized the sample size in Table 4. Most of the studies showed relatively small sample size in each age group.

Table 4. Participant size by chronological age in the included studies

Age	Kramer 2014		Ekizoglu 2016		Dedouit 2012		Ekizoglu 2015	
	M (n)	F (n)	M (n)	F (n)	M (n)	F (n)	M (n)	F (n)
8	-	-	-	-	-	-	2	2
9	-	-	-	-	-	-	2	1
10	3	2	5	5	5	8	2	3
11	0	4	10	5	3	4	3	5
12	6	5	10	6	8	5	7	4
13	5	6	10	9	7	12	6	5
14	6	5	23	10	8	6	6	2
15	9	2	16	15	8	11	9	3
16	8	9	26	14	10	5	7	4
17	7	6	10	6	6	6	6	3
18	7	5	11	3	6	7	3	5
19	7	6	3	4	5	11	9	3
20	9	4	5	5	7	3	2	4
21	13	9	15	10	5	11	4	7
22	8	6	22	7	9	4	5	5
23	9	11	20	13	13	11	8	3
24	10	7	19	10	5	10	9	6
25	12	6	17	19	5	5	7	5
26	9	5	23	10	7	7	-	-
27	12	4	17	10	4	3	-	-
28	8	6	13	10	8	8	-	-
29	8	8	17	15	7	7	-	-
30	10	8	13	12	2	8	-	-
Total	166	124	305	198	138	152	97	70

In addition, we converted these numbers to percentage distribution in each chronological age for both genders (Figure 5). All the four studies showed somewhat uneven size of participants in each age group for both genders.

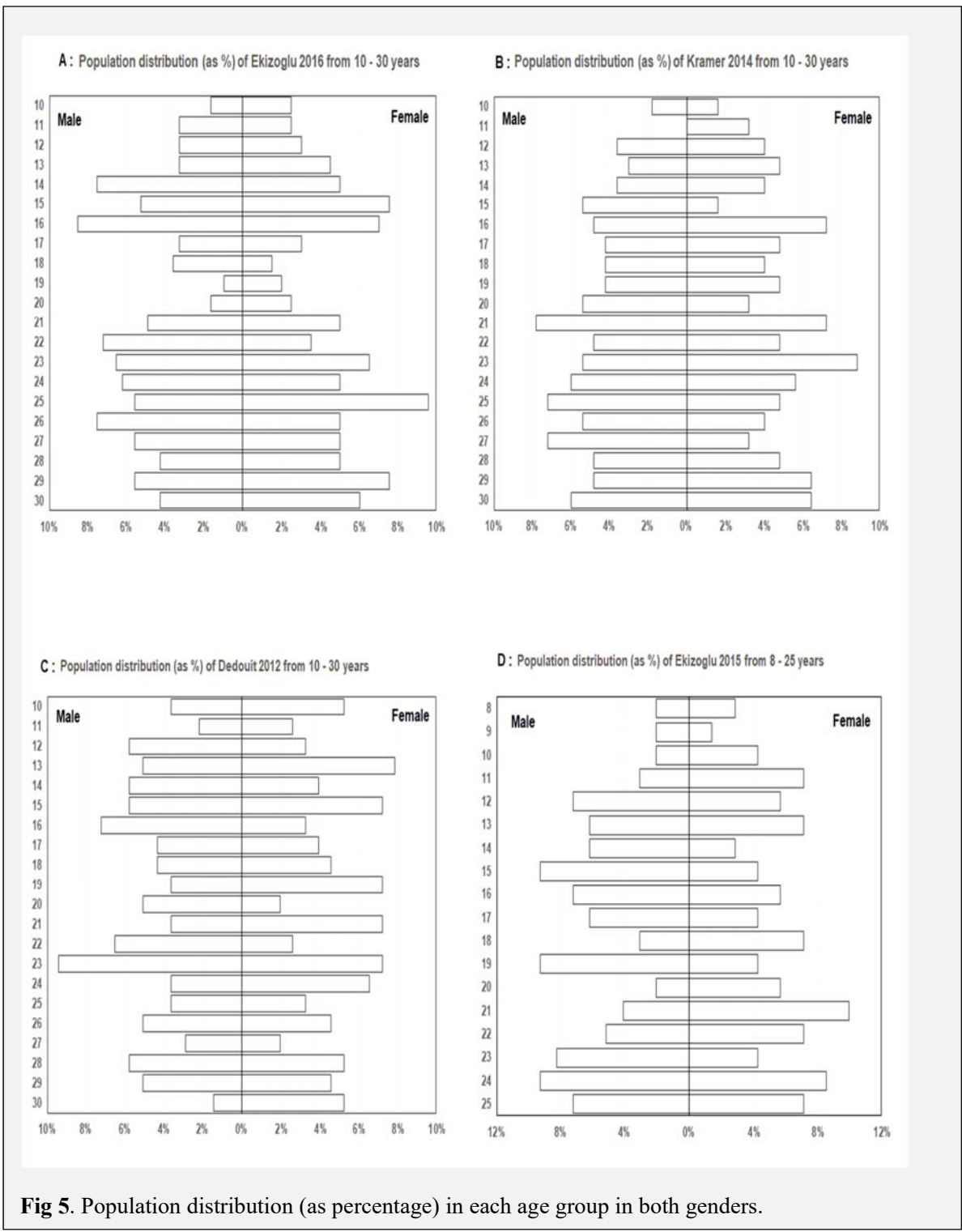


Fig 5. Population distribution (as percentage) in each age group in both genders.

Presentation of ossification stages

We first assessed whether it was appropriate to summarize the results of these studies in meta-analyses. QUADAS-2 checklist showed that all four studies were at high risk of being affected by age mimicry, and three studies had limited number of persons in most age groups. We believe that the lack of optimal sample size, and the potential risk of age mimicry would severely bias the pooled results and lead to inaccurate and unreliable estimates of chronological age in each knee ossification stage. Besides, we only found one paper on ankle ossification. Therefore, we concluded that it is not appropriate to compare and utilize the results in meta-analyses.

However, we present the findings graphically in the plots. First, we present the results related to ossification stages of the knee, beginning with distal femoral ossification with the MRI method described by Dedouit et al. (10) in two studies (Dedouit 2012 and Ekizoglu 2016), followed by the method of Kramer et al. [13] on distal femoral ossification in one study (Kramer 2014). In addition, we present the results on proximal tibial ossification with the method of Dedouit et al. in two studies (Dedouit 2012 and Ekizoglu 2016). In the end we present the findings of ankle ossification with the MRI method described by Saint-Martin et al (14) in one study (Ekizoglu 2015). Notably, we did not find any CT study on the ossification of knee and ankle.

Chronological age distribution and distal femoral ossification stages

Figure 6-15 illustrates the synthesized results of distal femoral ossification using the scoring method that was described by Dedouit et al. (10) in stage-specific and gender-specific manner. In the figure, CA is the mean chronological age with its 95% confidence interval, while size represents the total sample size in the corresponding stage from each study.

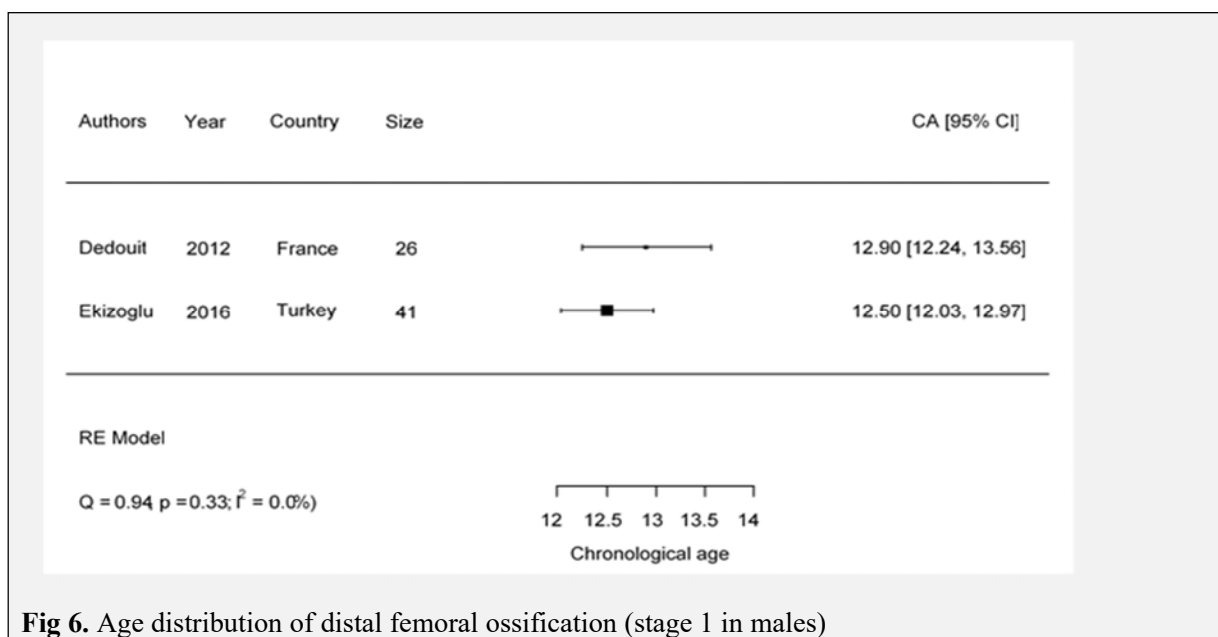


Fig 6. Age distribution of distal femoral ossification (stage 1 in males)

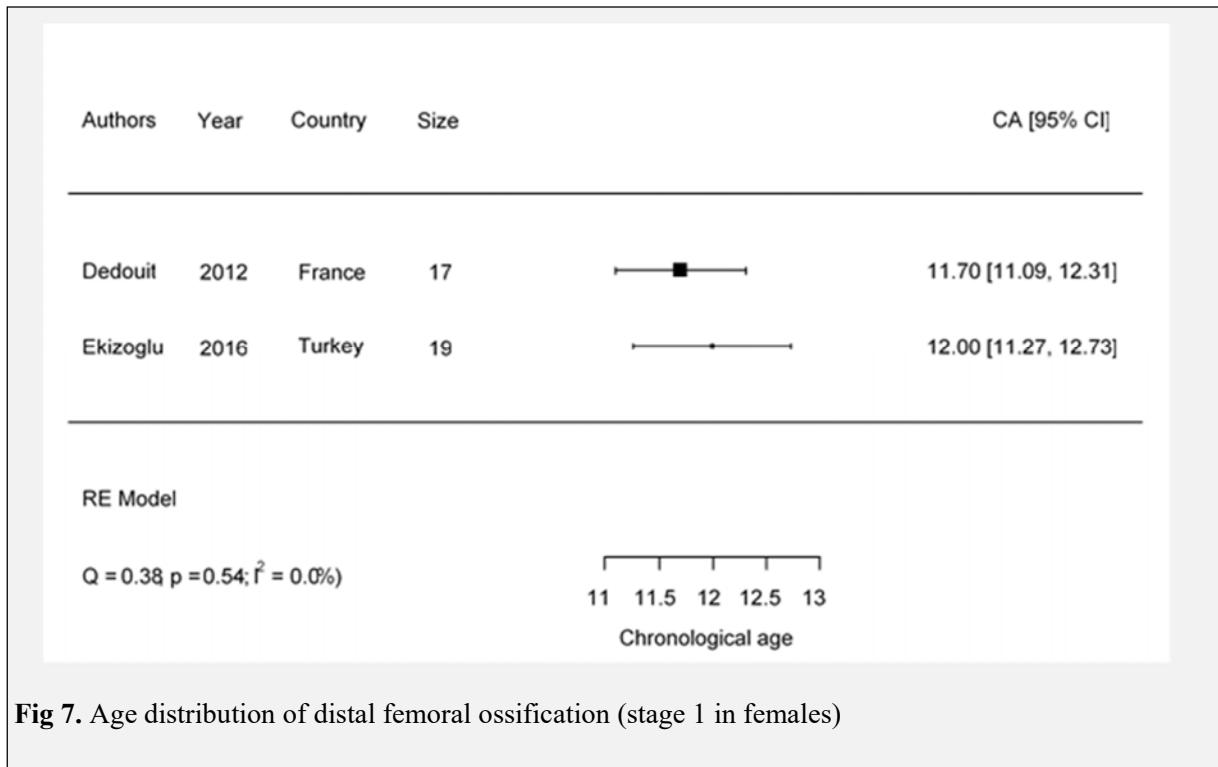


Fig 7. Age distribution of distal femoral ossification (stage 1 in females)

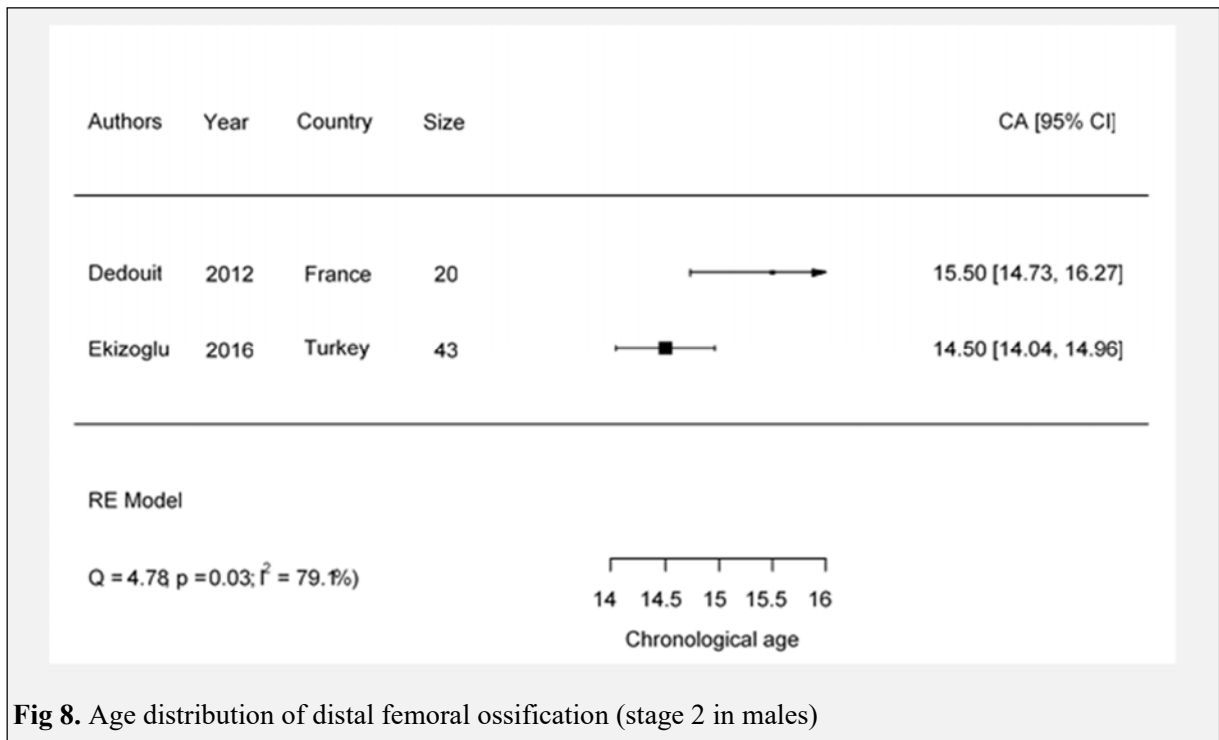


Fig 8. Age distribution of distal femoral ossification (stage 2 in males)

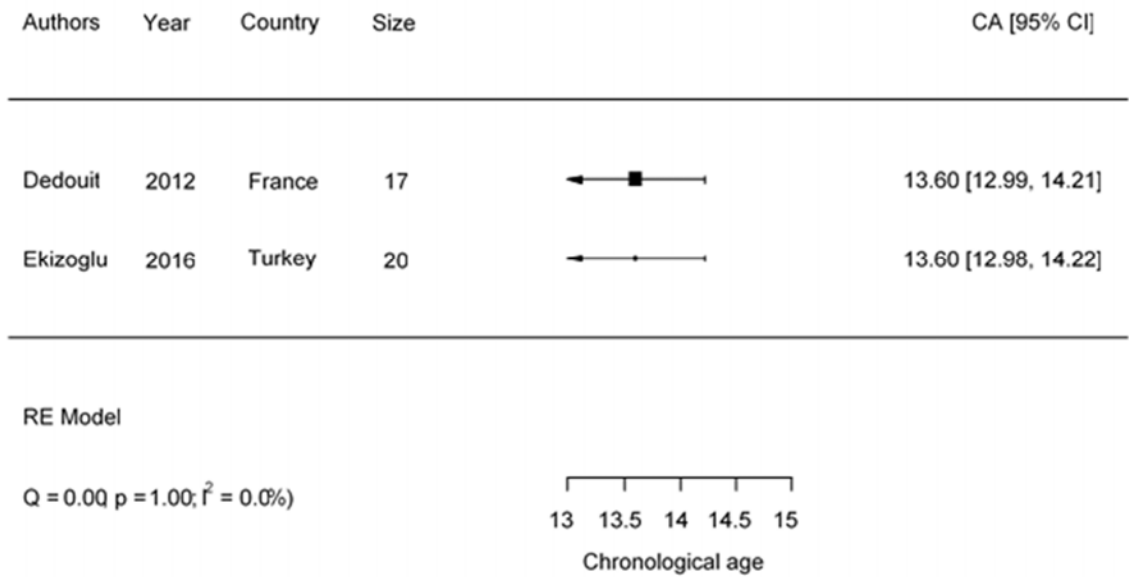


Fig 9. Age distribution of distal femoral ossification (stage 2 in females)

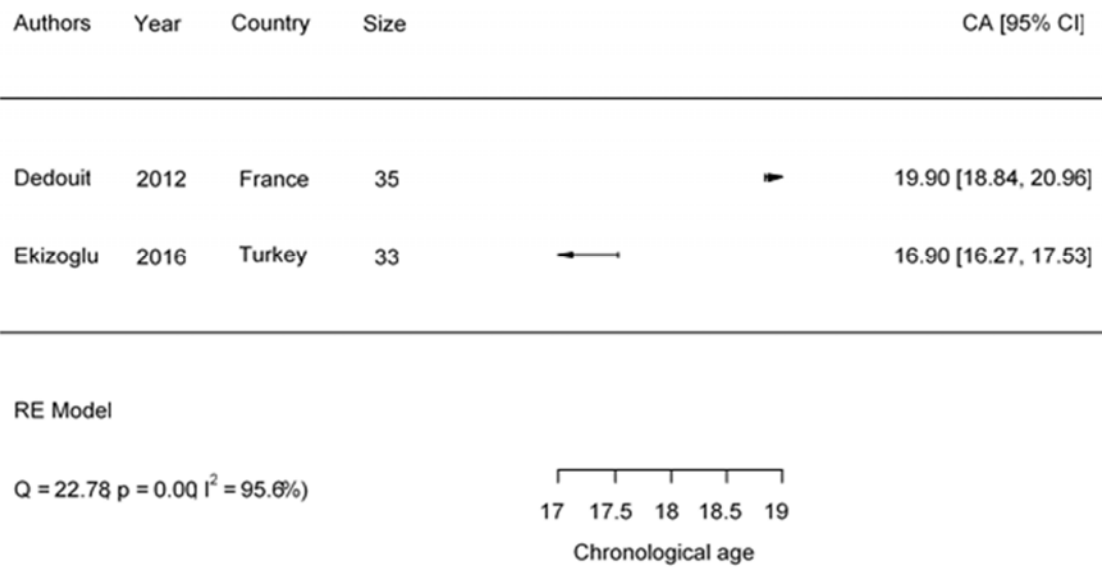


Fig 10. Age distribution of distal femoral ossification (stage 3 in males)

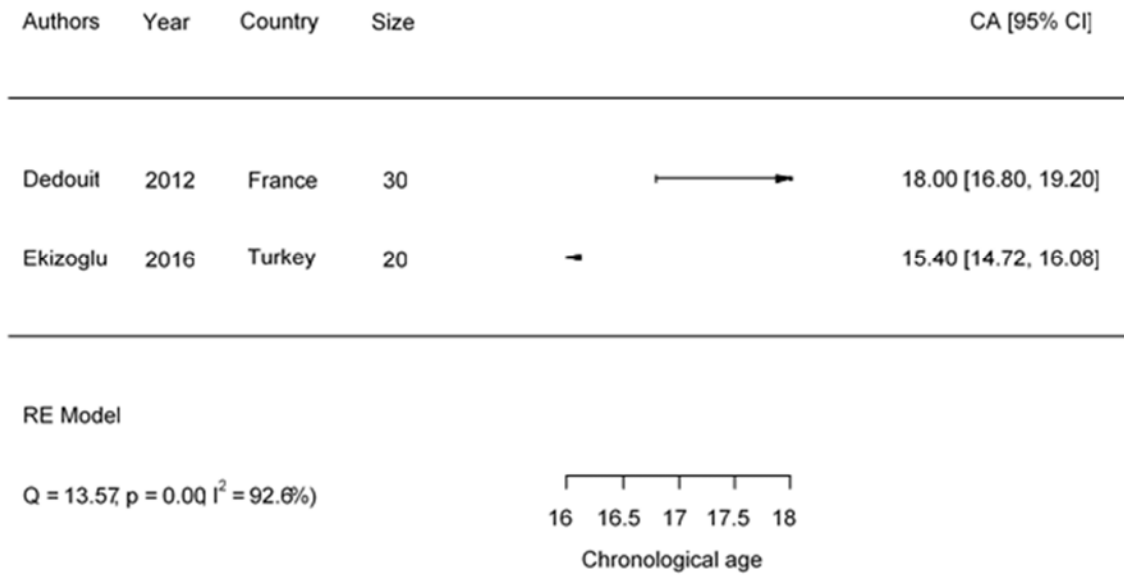


Fig 11. Age distribution of distal femoral ossification (stage 3 in females)

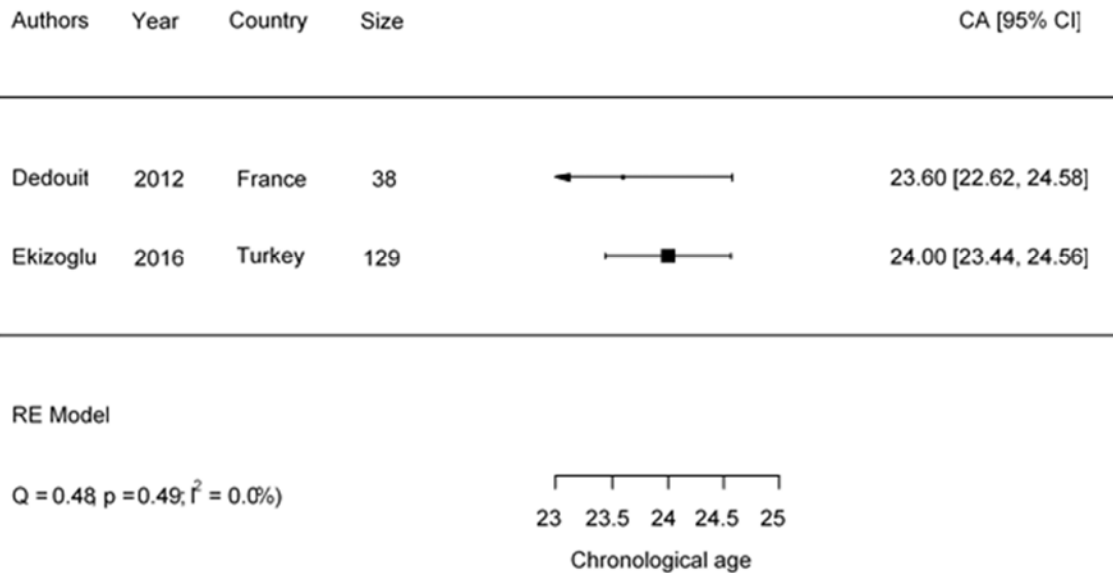


Fig 12. Age distribution of distal femoral ossification (stage 4 in males)

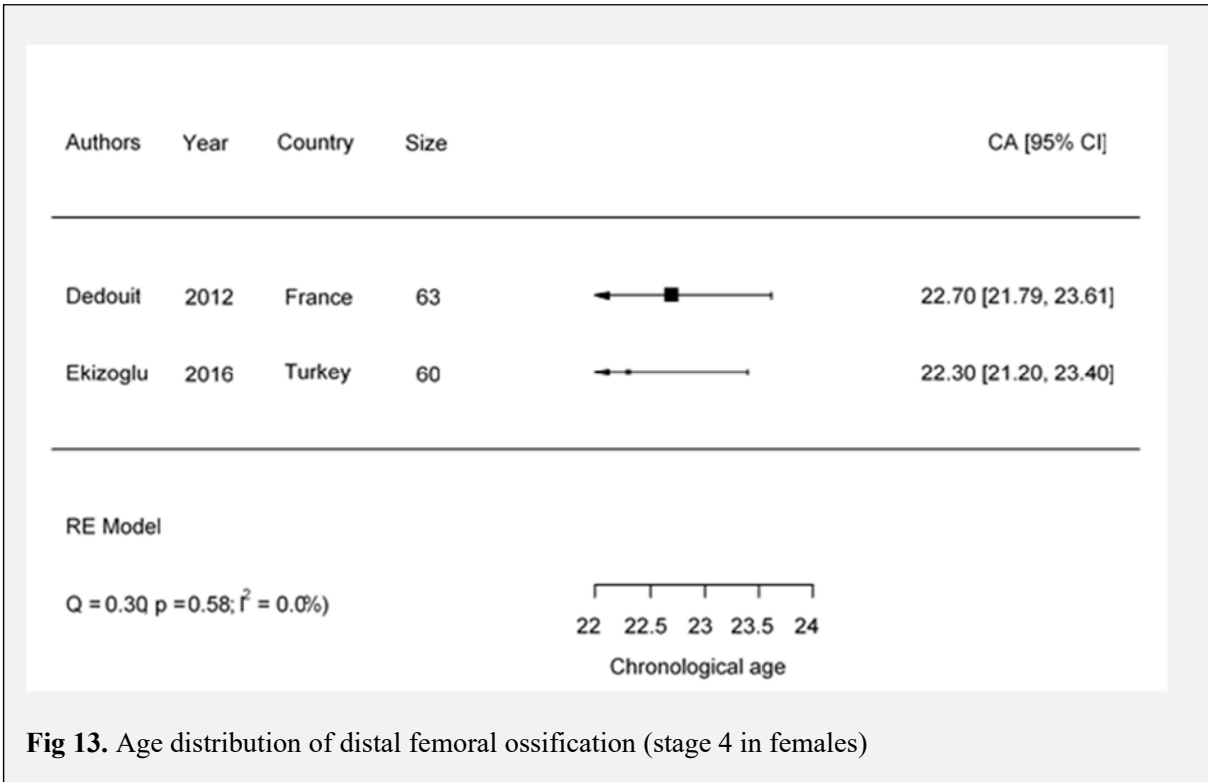


Fig 13. Age distribution of distal femoral ossification (stage 4 in females)

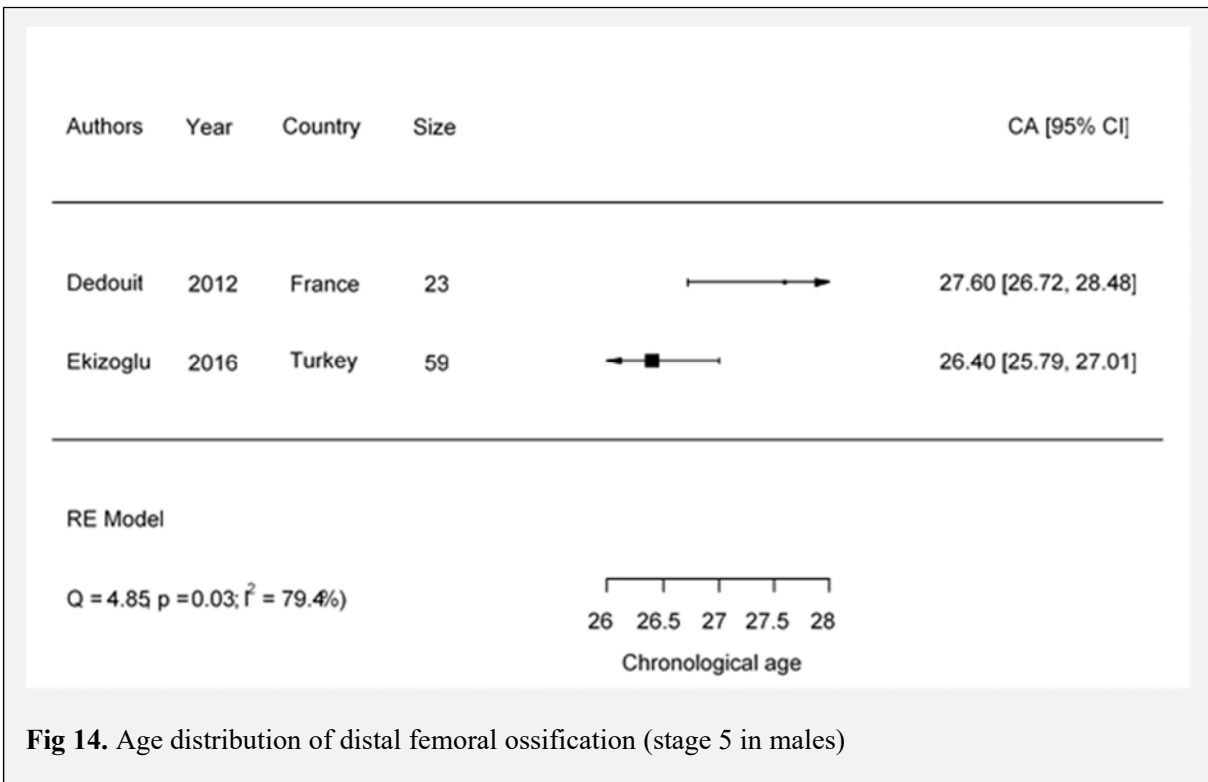


Fig 14. Age distribution of distal femoral ossification (stage 5 in males)

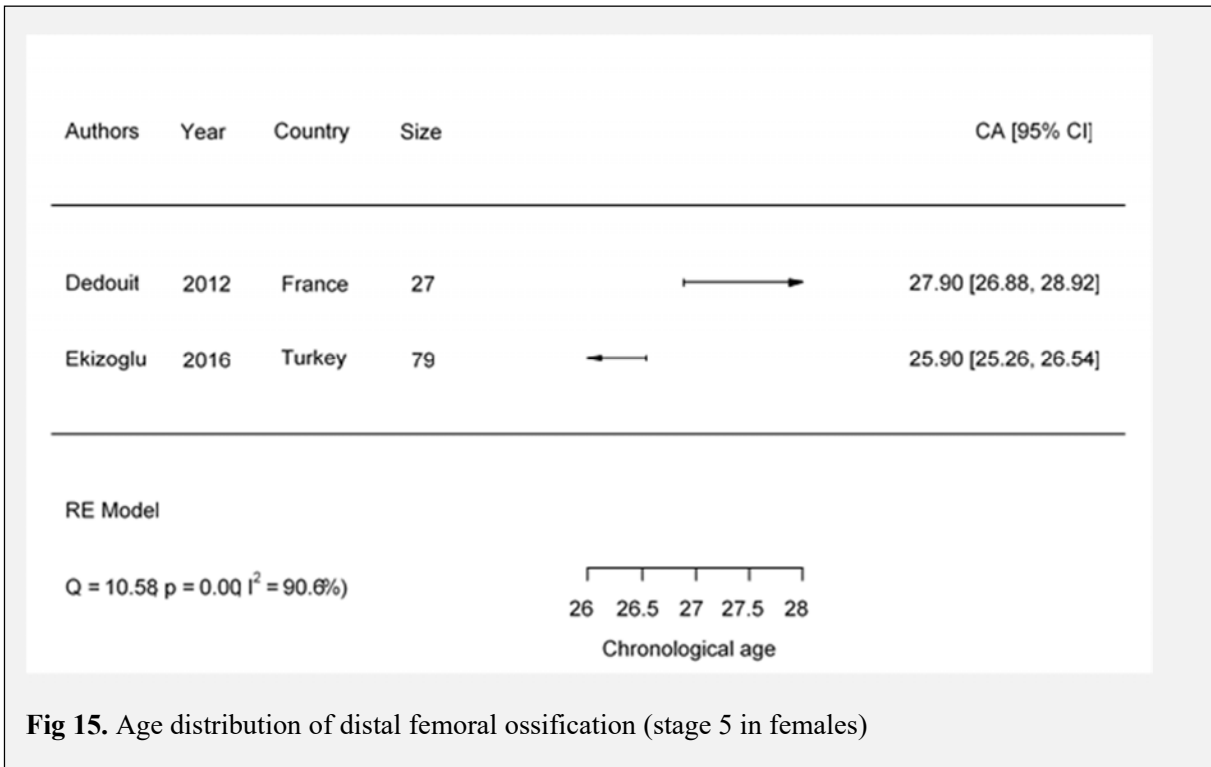


Fig 15. Age distribution of distal femoral ossification (stage 5 in females)

The results of distal femoral ossification in stage 1, 2 and 4 of the included studies showed consistent results on age distribution in both genders. However, in stage 3 and 5, we observed considerable heterogeneity between studies in both genders, participants of Dedouit 2012 in these stages are significantly older than Ekizoglu 2016. Notably, the number and the corresponding proportion of participants across age groups varied between the two studies.

In addition, we included one study that evaluated the ossification stages of distal femur with the MRI method described by Kramer et al (11). The results are shown in Table 5.

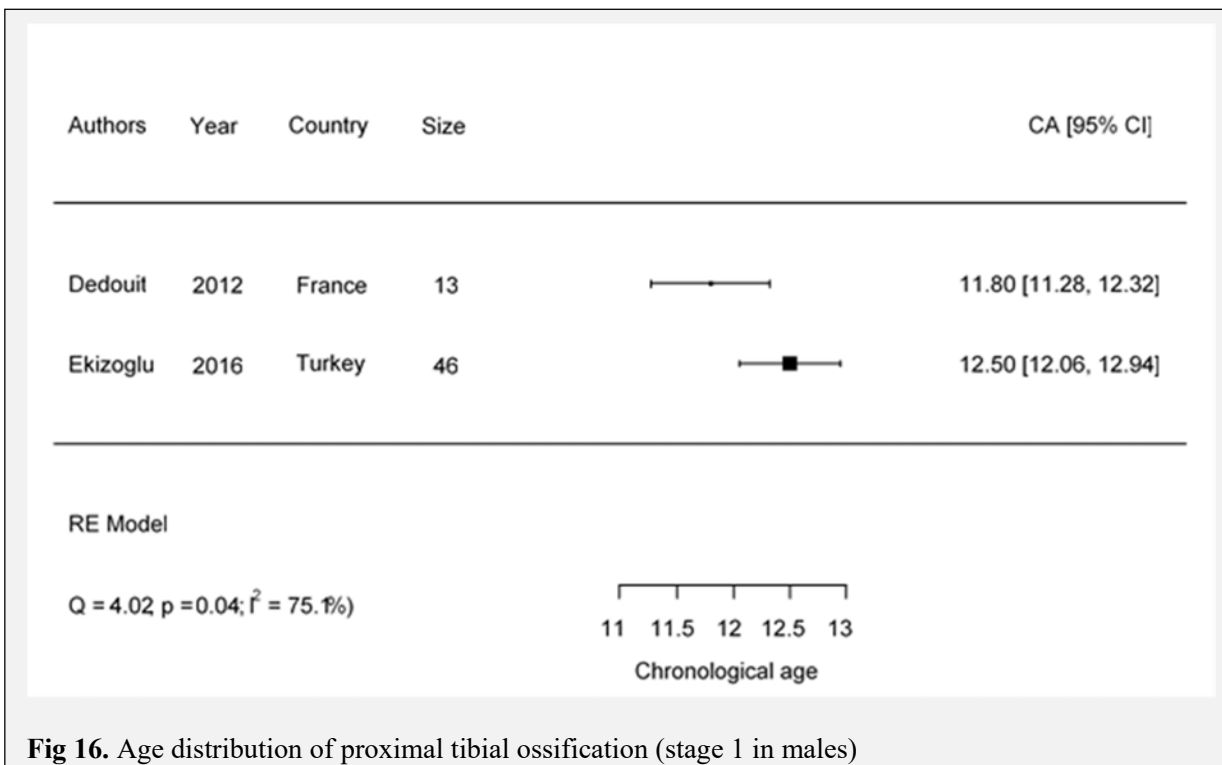
Table 5. Age distribution in each stage of distal femoral ossification (Kramer 2014)

	Males			Females		
	Sample size (n)	Mean age (years)	SD (years)	Sample size (n)	Mean age (years)	SD (years)
2c	8	12.30	1.70	5	11.80	1.40
3a	22	15.00	1.70	21	13.80	1.70
3b	2	15.10	0.10	-	-	-
3c	17	17.00	1.20	10	17.00	0.70
4	117	24.90	3.50	88	24.30	4.00

Participants in this study were from 10-30 years old. There were only 8 and 5 participants in stage 2c in males and females, with a mean of 12.30 (± 1.70) and 11.80 (± 1.40) years, respectively. There were 41 and 31 participants in stage 3 in male and female, respectively. Notably, the mean age of stage 3c in both male and female are much older than stage 3a. The majority of participants were in stage 4, with mean age of 24.90 (± 3.50) and 24.30 (± 4.00) years.

Chronological age distribution and proximal tibial ossification stages

Figure 16-25 illustrates the synthesized results of proximal tibial ossification in males and females using the scoring method that was described by Dedouit et al. (10). In the figure, CA is the chronological age with its 95% confidence interval, while size represents the total sample size in the corresponding stage from each study.



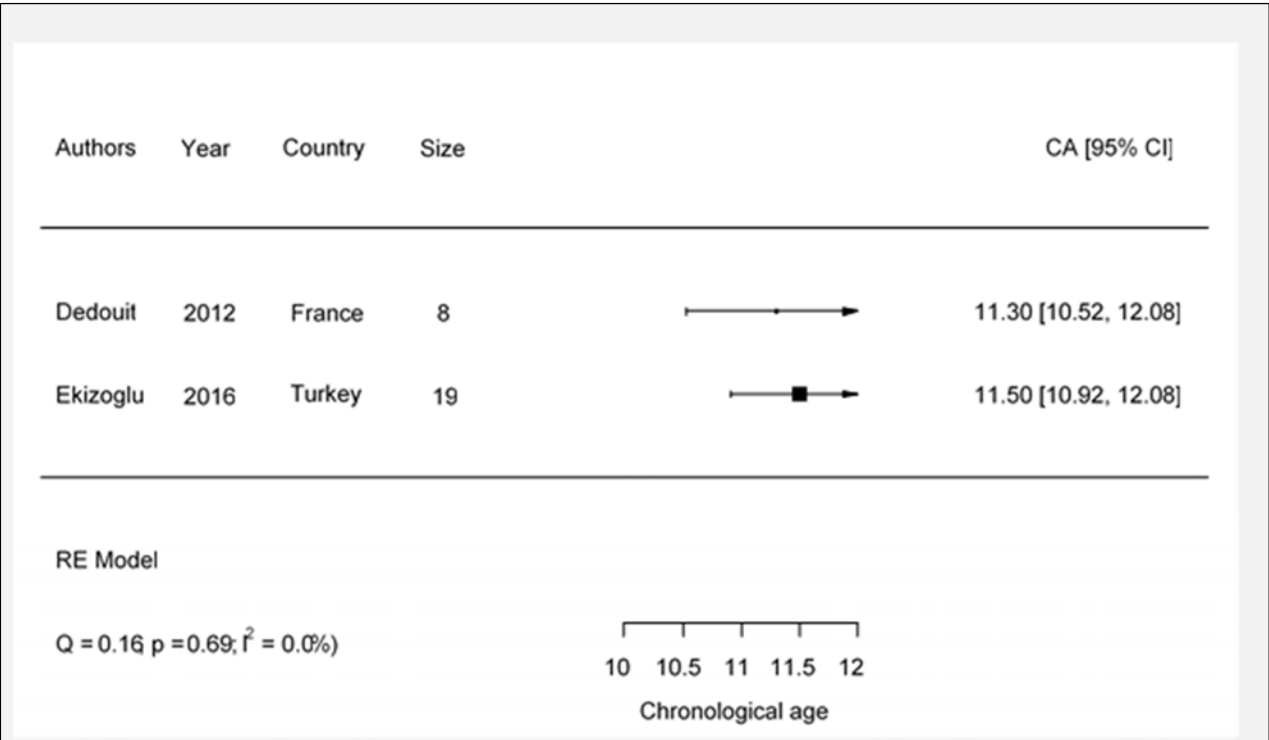


Fig 17. Age distribution of proximal tibial ossification (stage 1 in females)

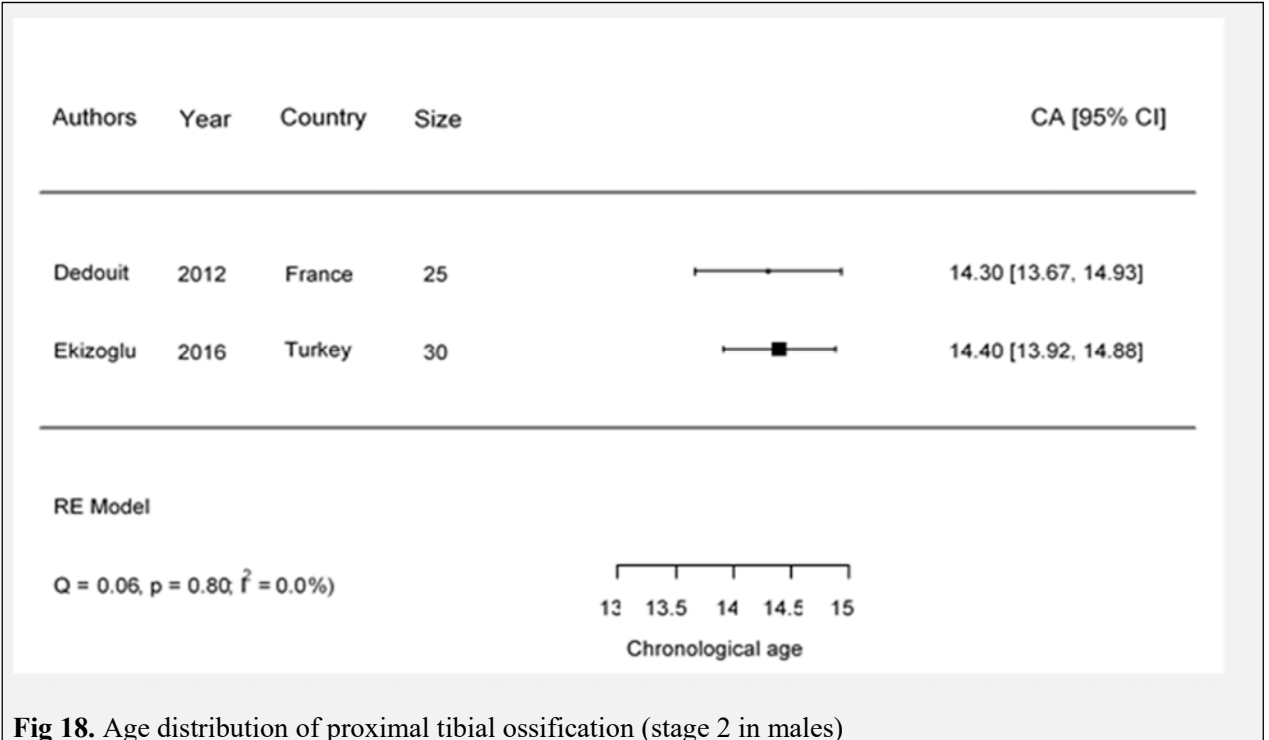


Fig 18. Age distribution of proximal tibial ossification (stage 2 in males)

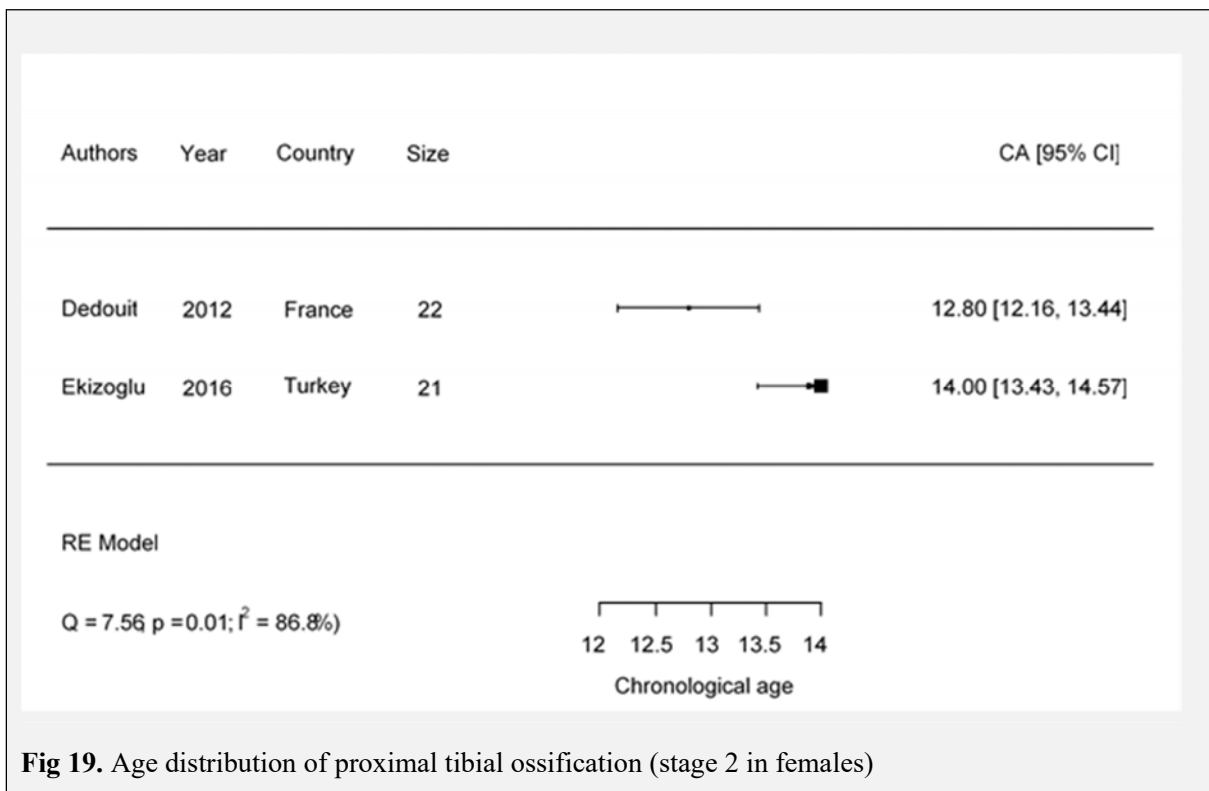


Fig 19. Age distribution of proximal tibial ossification (stage 2 in females)

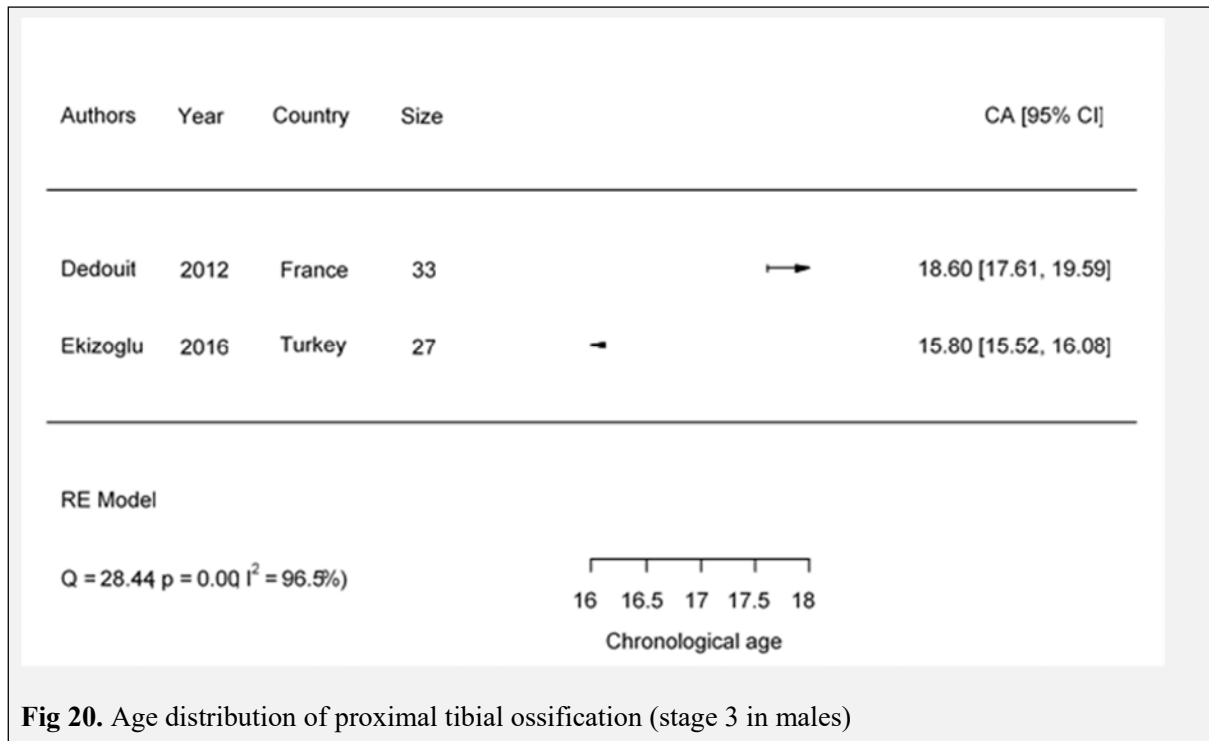


Fig 20. Age distribution of proximal tibial ossification (stage 3 in males)

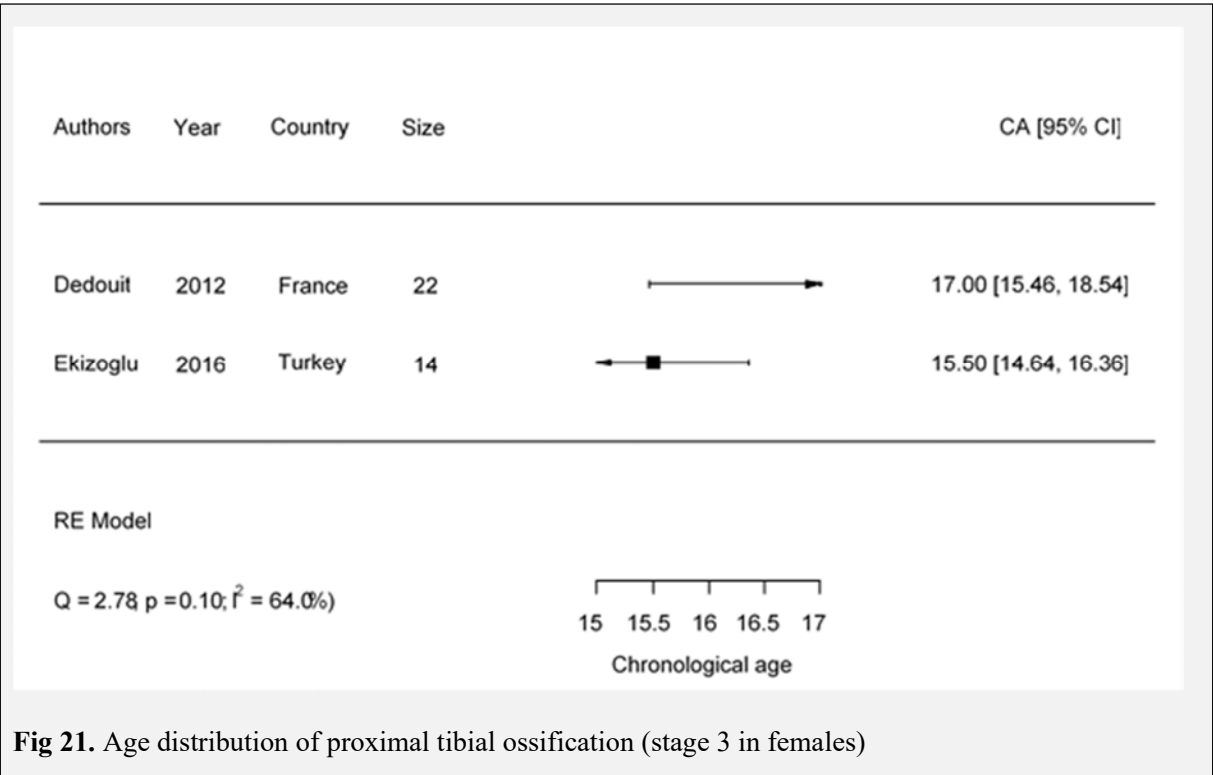


Fig 21. Age distribution of proximal tibial ossification (stage 3 in females)

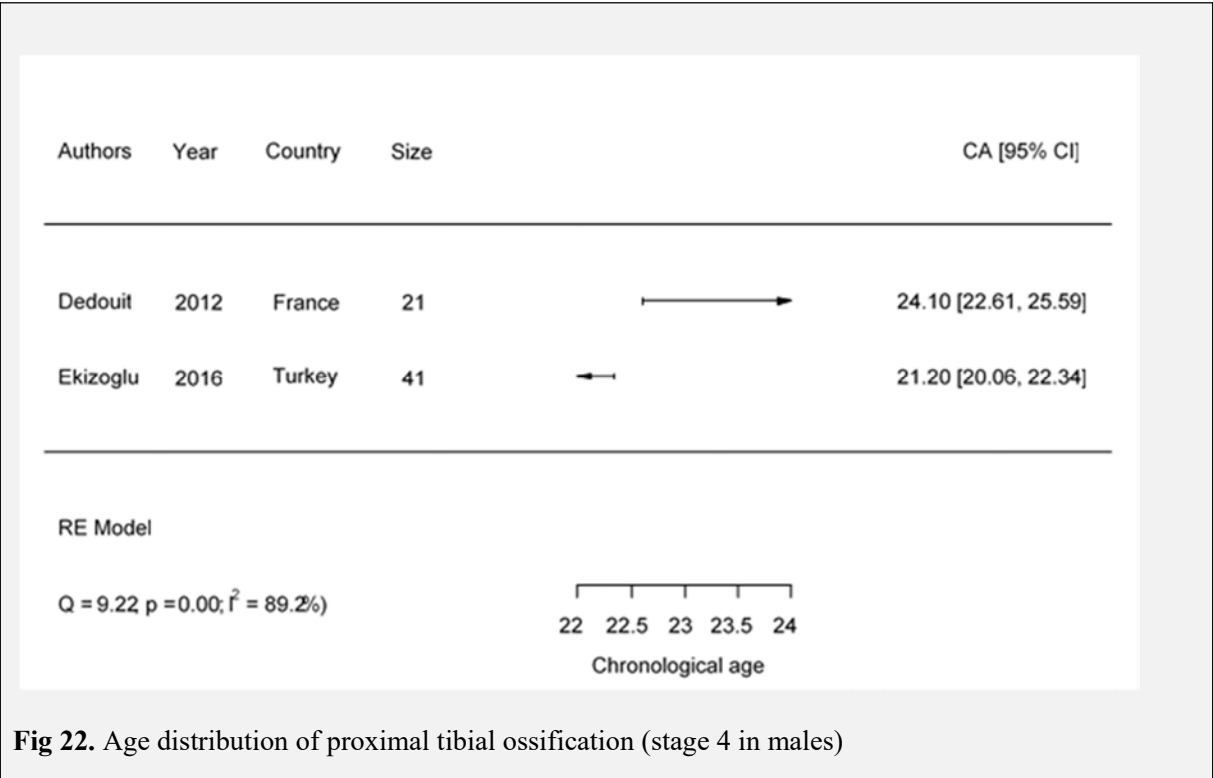


Fig 22. Age distribution of proximal tibial ossification (stage 4 in males)

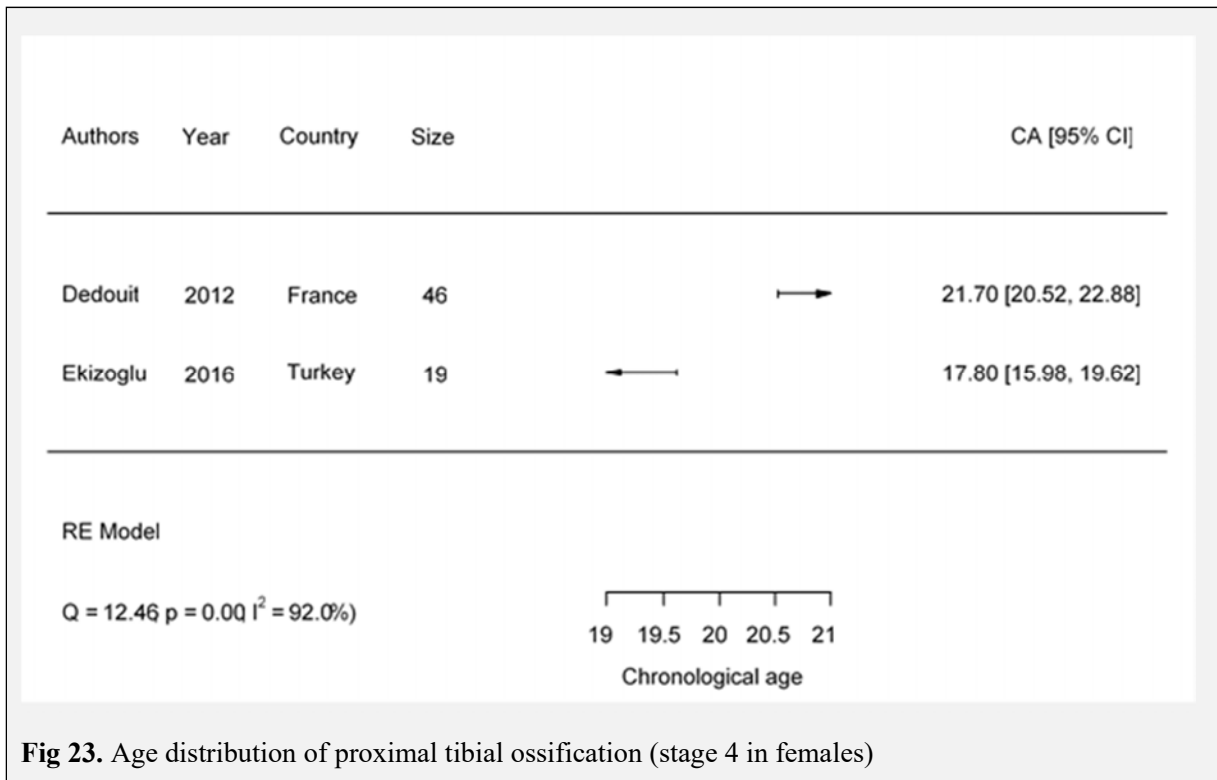


Fig 23. Age distribution of proximal tibial ossification (stage 4 in females)

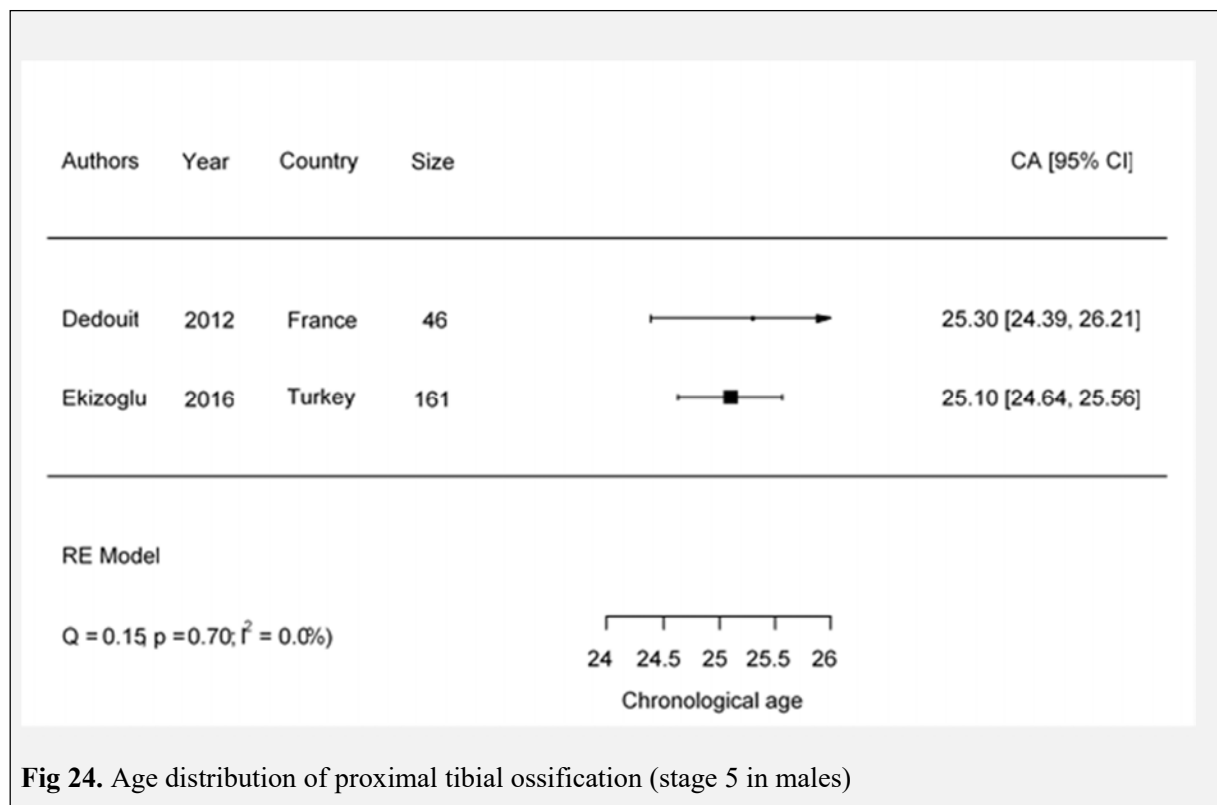


Fig 24. Age distribution of proximal tibial ossification (stage 5 in males)

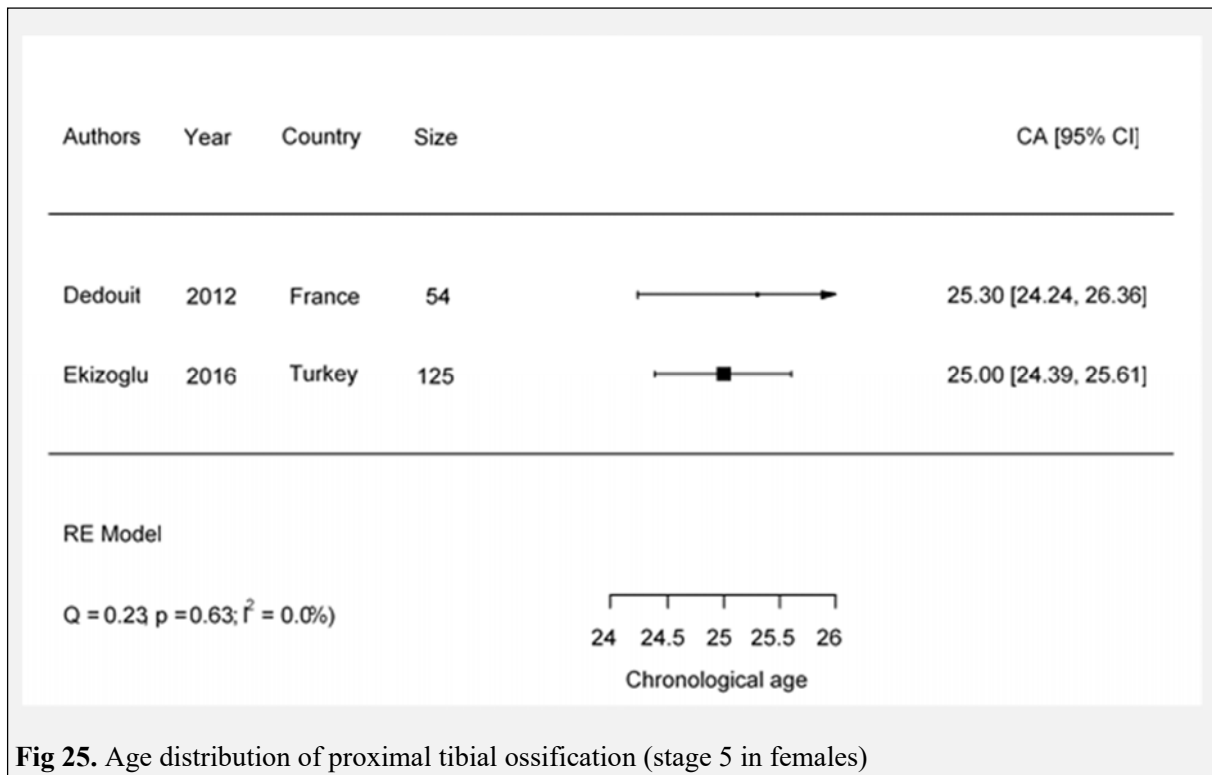


Fig 25. Age distribution of proximal tibial ossification (stage 5 in females)

Age distributions of most proximal tibial ossification stages varied considerably between included studies. Similar to the findings of distal femoral ossification, the mean age of participants from Dedouit et al (10) were generally higher than Ekizoglu et al (19) in most of the ossification stages. However, the age estimates of stage 5 in both genders were similar.

Chronological age distribution and ankle ossification stages

We found one study (13) that reported the chronological age distribution in each ankle ossification stage (Table 6) with a method described by Saint-Martin et al. (14). This method used a three-stage system, focusing on both distal tibial and calcaneus ossification.

Table 6. Age distribution in each stage of distal tibial and calcaneus ossification (Ekizoglu 2015)

	Distal tibia				Calcaneus			
	Male		Female		Male		Female	
	Sample size (n)	Mean age (±SD)	Sample size (n)	Mean age (±SD)	Sample size (n)	Mean age (±SD)	Sample size (n)	Mean age (±SD)
Stage 1	26	11.7±1.8	15	10.8±1.6	24	11.5±1.7	11	10.2±1.4
Stage 2	33	16.8±1.9	11	14.5±2.2	29	16.5±1.9	7	12.2±1.3
Stage 3	38	22.4±2.4	44	20.7±3.0	44	21.8±2.9	52	19.8±3.6

The differences of mean age between each stage in both genders are approximately 4-5 years. Notably, the standard deviation from stage 1 to stage 3 increased significantly, indicating a larger variation of chronological age in stage 3, comparing to stage 1 and 2.

Kappa agreement coefficients

We summarized the kappa coefficients for the intra- and inter-observer variabilities according to different methods on knee (Table 7) and ankle (Table 8) respectively.

Table 7. Cohen's kappa coefficients for the observer variability of different methods on knee

	Intra-observer		Inter-observer	
	Distal femur	Proximal tibia	Distal femur	Proximal tibia
Dedouit method				
Dedouit 2012	0.96	0.96	0.86	0.63
Ekizoglu 2016	0.92	0.96	0.84	0.89
Kramer method				
Kramer 2014	0.94	-	0.85	-

Table 8. Cohen's kappa coefficients for the observer variability on ankle

	Intra-observer		Inter-observer	
	Distal tibia	Calcaneum	Distal tibia	Calcaneum
Ekizoglu 2015	0.88	0.81	0.83	0.8

Most of the methods on knee ossification showed good observer accordance ($K > 0.80$), except the inter-observer coefficient on proximal tibia ($K = 0.63$). Similarly, the method on ankle ossification also showed promising kappa coefficients ($K > 0.80$).

Discussion

Summary and key findings

In this systematic review, we have summarized evidence on the chronological age distribution according to the ossification stages of knee and ankle by CT and MRI. Four MRI studies meeting the inclusion criteria were included. We did not find any eligible study using CT.

There were two MRI classification methods focusing on the knee ossification and one MRI classification method on the ankle ossification. All of the studies covered a relatively wide age spectrum, but were comprised of limited or uneven numbers of participants in each age group, leading to inaccurate estimation of age distribution. All methods showed good intra- and inter-observer agreements. However, since the number of included studies is so few, more studies are needed to validate the applicability of these methods for age estimation.

Quality of the results

In this review, we found two MRI methods that have assessed the ossification stages of the knee.

The first method was suggested by Dedouit et al. (10) and Ekizoglu 2016 et al. (19), which examines the staging for distal femoral and proximal tibial epiphyses separately. Both studies covered a relatively wide age spectrum, but with limited number of participants in each age group. We observed significant inconsistency of age distribution in both genders in stage II, III and IV of distal femoral and proximal tibial ossification, which suggests large variation of results in the included studies. Since both studies have low or unclear risk of bias in the “Patient selection” domain of QUADAS-2 checklist, general patient selection process might not differentiate significantly. Heterogeneity may come from the following sources: age mimicry bias, variation due to limited number of participants, and/or systematic difference between studies (MRI settings, different maturation rates between populations, systematic imaging reading difference etc.).

As shown in Table 2, none of the included studies adopted the same MRI settings, which might be another source for the discrepancy of findings in the current systematic review (11). Further, Ekizoglu et al. (19) compared their findings with the findings from Dedouit et al. (10), in which both studies used the same sequence (T2) but different Tesla values (3.0 T versus 1.5 T). The findings indicate adequate agreement

between the two studies in terms of the first ages at which certain ossification stages were evident. However, the mean age (with SD) varied considerably in at least some stages of both proximal tibial and distal femoral epiphysis between the two studies, which implies the existence of other sources of bias.

In fact, we consider age mimicry bias and small sample size as the major limitations of most of the included studies. Both problems may cause large variation on age distribution in all the ossification stages, and lead to inaccurate and inconsistent results when presenting data together. Notably, all four included studies were aiming to describe the age distribution but using different grading system or MRI settings, and statistical power analyses were not mentioned in these studies.

The second method for evaluating the ossification staging of knee was suggested by Kramer and his colleagues (11). Kramer et al. (11) modified a five-stage classification method that has been used on medial clavicle epiphysis (5, 12). Unlike the method suggested by Dedouit et al. (10), this method does not require assessment of proximal tibial ossification stages, only focusing on distal femoral ossification. Furthermore, this method has sub-categories on both stage 2 and 3 of distal femoral ossification, which might provide more relevant information on 18 years threshold. Unfortunately, we only found one paper focusing on this MRI method, therefore its validity should be explored further in larger samples.

We included one study assessing the validity of ankle ossification on age assessment (13), which used a three-stage scoring system on distal tibial and calcaneus ossification. The sample size in every chronological year was relatively small (2-9 persons every chronological year), leading to high risk of age mimicry bias and large variation of chronological age distribution. In particular, the standard deviation of mean age in stage 3 for both genders are larger than the other two stages, indicating higher risk of imprecision on chronological age estimation in stage 3.

Another issue that is worth to mention is the “end-stage problem”. This is an inevitable topic in age estimation studies, and occurs when the biological development reaches its end stage. The mean age and variance of the end stage will be highly influenced by the upper age limit of the included population. Therefore, if the studies extended their population age spectrum, the mean age of the end stage would be susceptible to this change. This easily overlooked issue has been mentioned in several studies (20-22) and is discussed in more detail in our consortium papers on age estimation with hand-wrist and teeth (1, 2). Notably, all of the included studies on knee ossification included populations with the same age spectrum (10-30 years old). Hence the upper age limit might not affect the mean age significantly across the included studies.

Although the included studies suffer from limited number of participants and possibly age mimicry bias, the results of kappa coefficients of all three methods on knee and ankle seemed to have good intra- and inter-observer reliabilities. It is, however, necessary to test the agreement of those methods in larger samples.

Future perspective

In our opinion, age mimicry is an important source of bias in age assessment studies, which also makes it less sensible to summarize individual studies that did not take this issue into account (for studies presenting mean age of a certain stage). Future studies are warranted with even number of participants in each age, covering a wide age spectrum and with a large sample size. An alternative approach is to collect the original research data from current studies, re-distribute and analyze them with other strategies for more reliable results.

In practice, forensic experts often receive CT or MRI images to evaluate which ossification stage this image belonging to, then estimate individual's age accordingly. Therefore, it is crucial to have a reference age distribution for maturation stages of knee and ankle. Several age estimation studies using medial clavicle ossification stages (22, 23) have proposed transition analysis as a promising analytical technique for obtaining a reference age distribution. This approach estimates probability-based age of a population according to Bayes' theorem, which is known to be less sensitive to age mimicry bias (24). This method might also be useful for age estimation with knee and ankle ossification.

Besides, all four included analysis were carried out in Europe. It is unclear if there is systematic difference on knee and ankle ossification across remote regions and ethnicities, future studies representing other regions or populations are needed for a more comprehensive understanding of maturation development of knee and ankle ossification.

Strengths and limitations

The strength of this systematic review is the systematic and transparent approach that we have used to review the question. We have implemented systematic literature search in many electronic databases, with clear inclusion and exclusion criteria. Two of the authors have independently considered each reference according to these criteria, and assessed the quality of the included studies. These independent assessments reduce the risk of bias. Although we have conducted a thorough literature search, potential relevant studies might not have been identified. A built-in weakness with systematic reviews is that they may become outdated when new studies are published. This systematic review is up-to-date as of April 2017.

As the purpose of this systematic review was to evaluate age assessment in the living, we excluded studies of corpses, and this might be regarded as a limitation of our systematic review. However, we argue that the differences between the two groups are not fully understood (e.g. movement artifacts in the study of living persons), which is the reason for our choice.

The current systematic review was one of the several reports undertaken to evaluate different age estimation methods in order to improve the Norwegian system for age assessment of minor asylum seekers. The assignment given was to estimate a probable age, and the probability of being 16 or 18 years. Hence, we chose to focus on the most probable age range (mean age and SD) given the observed

development stages. A different approach is to focus on the minimum age (ever observed for an individual with a certain observed development stage of e.g. the distal femur), which is not sensitive to age mimicry bias. With this approach, one may reach a high degree of certainty as to whether the individual is above eighteen. However, the minimum age concept is vulnerable in the sense that one or a few outliers might move the estimate substantially. Also, even though the concept will lead to few children being assessed as adults, a great deal of adults will be assessed as children.

Conclusion

In this systematic review, we have summarized evidence that describes the age distribution according to knee and ankle ossification stages by CT and MRI. We did not find any CT study about age assessment of ossification of the knee and ankle. Four MRI studies were included, documenting two methods for assessing the knee ossification, and one method for ankle ossification. All three methods showed good intra- and inter-observer reliability. However, due to the limited number of participants in these studies and age mimicry bias, we are unable to assess the agreement of those methods with chronological age. To get more reliable results, future studies need to have sufficient number of participants, uniform age distribution, and an age spectrum covering the possible chronological ages of every developmental stage analyzed.

References

1. Dahlberg PS, Mosdøl A, Ding KY, Bleka Ø, Rolseth V, Skjerven-Martinsen M, Straumann GH, Delaveris GJM, Vist GE. Samsvar mellom kronologisk alder og skjelettalder basert på Greulich og Pyle - atlaset for aldersestimering: en systematisk oversikt [Agreement between chronological age and bone age based on the Greulich and Pyle-atlas for age estimation: a systematic review]. Rapport 2017. Oslo: Folkehelseinstituttet, 2017. <https://www.fhi.no/publ/2017/samsvar-mellom-kronologisk-alder-og-skjelettalder-basert-pa-greulich--pyle-/>.
2. Rolseth V, Mosdøl A, Dahlberg PS, Ding KY, Bleka Ø, Skjerven-Martinsen M, Straumann GH, Delaveris GJM, Vist GE. Demirjians utviklingsstadier på visdomstenner for estimering av kronologisk alder: en systematisk oversikt. [Demirjian's development stages on wisdom teeth for estimation of chronological age: a systematic review.] Rapport 2017. Oslo: Folkehelseinstituttet, 2017. <https://www.fhi.no/publ/2017/demirjians-utviklingsstadier-pa-visdomstenner-for-estimering-av-kronologisk/>
3. Bassed RB, Drummer OH, Briggs C, Valenzuela A. Age estimation and the medial clavicular epiphysis: analysis of the age of majority in an Australian population using computed tomography. *Forensic science, medicine, and pathology*. 2011;7:148-54.
4. Schulze D, Rother U, Fuhrmann A, Richel S, Faulmann G, Heiland M. Correlation of age and ossification of the medial clavicular epiphysis using computed tomography. *Forensic Science International*. 2006;158:184-9.
5. Schmeling A, Schulz R, Reisinger W, Muhler M, Wernecke KD, Geserick G. Studies on the time frame for ossification of the medial clavicular epiphyseal cartilage in conventional radiography. *International journal of legal medicine*. 2004;118:5-8.
6. Schulz R, Muhler M, Mutze S, Schmidt S, Reisinger W, Schmeling A. Studies on the time frame for ossification of the medial epiphysis of the clavicle as revealed by CT scans. *International journal of legal medicine*. 2005;119:142-5.
7. Soegiharto BM, Cunningham SJ, Moles DR. Skeletal maturation in Indonesian and white children assessed with hand-wrist and cervical vertebrae methods. *American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics*. 2008;134:217-26.
8. Cole TJ, Rousham EK, Hawley NL, Cameron N, Norris SA, Pettifor JM. Ethnic and sex differences in skeletal maturation among the Birth to Twenty cohort in South Africa. *Archives of disease in childhood*. 2015;100:138-43.
9. Cameriere R, Cingolani M, Giuliadori A, De Luca S, Ferrante L. Radiographic analysis of epiphyseal fusion at knee joint to assess likelihood of having attained 18 years of age. *International journal of legal medicine*. 2012;126:889-99.
10. Dedouit F, Auriol J, Rousseau H, Rouge D, Crubezy E, Telmon N. Age assessment by magnetic resonance imaging of the knee: a preliminary study. *Forensic Sci Int*. 2012;217:232.e1-7.

11. Kramer JA, Schmidt S, Jurgens KU, Lentschig M, Schmeling A, Vieth V. Forensic age estimation in living individuals using 3.0 T MRI of the distal femur. *International journal of legal medicine*. 2014;128:509-14.
12. Kellinghaus M, Schulz R, Vieth V, Schmidt S, Pfeiffer H, Schmeling A. Enhanced possibilities to make statements on the ossification status of the medial clavicular epiphysis using an amplified staging scheme in evaluating thin-slice CT scans. *International journal of legal medicine*. 2010;124:321-5.
13. Ekizoglu O, Hocaoglu E, Can IO, Inci E, Aksoy S, Bilgili MG. Magnetic resonance imaging of distal tibia and calcaneus for forensic age estimation in living individuals. *International journal of legal medicine*. 2015;129:825-31.
14. Saint-Martin P, Rerolle C, Dedouit F, Bouilleau L, Rousseau H, Rouge D, Telmon N. Age estimation by magnetic resonance imaging of the distal tibial epiphysis and the calcaneum. *International journal of legal medicine*. 2013;127:1023-30.
15. helsetjenesten Nkf. Slik oppsummerer vi forskning. Håndbok for Nasjonalt kunnskapssenter for helsetjenesten. 3.2. reviderte utg. . Oslo: Nasjonalt kunnskapssenter for helsetjenesten. 2013.
16. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Systematic reviews*. 2016;5:210.
17. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. 2010. 2010;36:48.
18. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine*. 2011;155:529-36.
19. Ekizoglu O, Hocaoglu E, Inci E, Can IO, Aksoy S, Kazimoglu C. Forensic age estimation via 3-T magnetic resonance imaging of ossification of the proximal tibial and distal femoral epiphyses: Use of a T2-weighted fast spin-echo technique. *Forensic Sci Int*. 2016;260:102.e1-e7.
20. Olze A, Pynn BR, Kraul V, Schulz R, Heinecke A, Pfeiffer H, Schmeling A. Studies on the chronology of third molar mineralization in First Nations people of Canada. *International journal of legal medicine*. 2010;124:433-7.
21. Roberts GJ, McDonald F, Andiappan M, Lucas VS. Dental Age Estimation (DAE): Data management for tooth development stages including the third molar. Appropriate censoring of Stage H, the final stage of tooth development. *Journal of forensic and legal medicine*. 2015;36:177-84.
22. Langley-Shirley N, Jantz RL. A Bayesian approach to age estimation in modern Americans from the clavicle. *Journal of forensic sciences*. 2010;55:571-83.
23. Franklin D, Flavel A. CT evaluation of timing for ossification of the medial clavicular epiphysis in a contemporary Western Australian population. *International journal of legal medicine*. 2015;129:583-94.
24. Boldsen JL, Milner GR, Konigsberg LW, Wood JW. Transition analysis: a new method for estimating age from skeletons. In: Vaupel JW, Hoppa RD, editors. *Paleodemography: Age Distributions from Skeletal Samples*. Cambridge Studies in Biological and Evolutionary Anthropology. Cambridge: Cambridge University Press; 2002. p. 73-106.

Appendices

Appendix 1. Literature search strategy

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present

Search date: 2016-05-19

1. Age Determination by Teeth/ (1422)
2. Age Determination by Skeleton/ (3937)
3. (age adj3 (determinat* or estimat* or assess*)).ti. (2851)
4. ((forensic or radiological) adj age).ti,ab. (158)
5. ((age or maturation or mature or ossification) adj5 (determinat* or estimat* or assess* or examinat*)).ti,ab. (41703)
6. (hand\$1 or wrist\$ or carpal or metacarpal or metacarpus or dental or teeth or tooth or third molar* or clavicle* clavícula* or collar bone* or femur or tibia* or fibula* or knee or knees or foot or feet or ankle or ankles).ti,ab. (904235)
7. (MRI or MR imag* or magnetic resonance imag* or ct scan* or cat scan* or (comput* adj2 tomograp*) or roentgen or x-ray* or xray* or radiolog* or radiograp*).ti,ab. (1032026)
8. 5 and 6 and 7 (1297)
9. 1 or 2 or 3 or 4 or 8 (7491)
10. exp Animals/ (20185560)
11. Humans/ (15941900)
12. 10 not (10 and 11) (4243660)
13. 9 not 12 (7007)
14. (greulich adj2 pyle).ti,ab. (238)

15. (tanner adj2 whitehouse).ti,ab. (246)
16. demirjian.ti,ab. (218)
17. haavikko.ti,ab. (20)
18. kullman.ti,ab. (6)
19. nortje.ti,ab. (5)
20. liversidge.ti,ab. (10)
21. kvaal.ti,ab. (13)
22. or/14-21 (674)
23. 13 or 22 (7178)

Database: Embase 1974 to 2016 May 18

Search date: 2016-05-19

1. age determination/ (5176)
2. (age adj3 (determinat* or estimat* or assess*)).ti. (3291)
3. ((forensic or radiological) adj age).ti,ab. (199)
4. ((age or maturation or mature or ossification) adj5 (determinat* or estimat* or assess* or examinat*)).ti,ab. (57474)
5. (hand\$1 or wrist\$ or carpal or metacarpal or metacarpus or dental or teeth or tooth or third molar* or clavicle* clavicula* or collar bone* or femur or tibia* or fibula* or knee or knees or foot or feet or ankle or ankles).ti,ab. (1087091)
6. (MRI or MR imag* or magnetic resonance imag* or ct scan* or cat scan* or (comput* adj2 tomograp*) or roentgen or x-ray* or xray* or radiolog* or radiograp*).ti,ab. (1334461)
7. 4 and 5 and 6 (1656)
8. 1 or 2 or 3 or 7 (8121)
9. exp animals/ or exp invertebrate/ or animal experiment/ or animal model/ or animal tissue/ or animal cell/ or nonhuman/ (23089391)
10. human/ or normal human/ or human cell/ (17222575)
11. 9 not (9 and 10) (5913580)
12. 8 not 11 (7315)

13. (greulich adj2 pyle).ti,ab. (338)
14. (tanner adj2 whitehouse).ti,ab. (279)
15. demirjian.ti,ab. (208)
16. haavikko.ti,ab. (19)
17. kullman.ti,ab. (7)
18. nortje.ti,ab. (4)
19. liversidge.ti,ab. (18)
20. kvaal.ti,ab. (11)
21. or/13-20 (794)
22. 12 or 21 (7692)

Database: Central

Search date: 2016-05-19

- #1 MeSH descriptor: [Age Determination by Skeleton] explode all trees (99)
- #2 MeSH descriptor: [Age Determination by Teeth] explode all trees (5)
- #3 (age near/3 (determinat* or estimat* or assess*)):ti (30)
- #4 ((forensic or radiological) next age) (0)
- #5 ((age or maturation or mature or ossification) near/5 (determinat* or estimat* or assess* or examin*)) (3474)
- #6 (hand or hands or wrist or wrists or carpal or metacarpal or metacarpus or dental or teeth or tooth or third molar* or clavicle* clavícula* or collar bone* or femur or tibia* or fibula* or knee or knees or foot or feet or ankle or ankles) (78361)
- #7 (MRI or (MR next imag*) or (magnetic next resonance next imag*) or ct-scan* or cat-scan* or (comput* near/2 tomograp*) or roentgen or x-ray* or xray* or radiolog* or radiograp*) (52159)
- #8 #5 and #6 and #7 (236)
- #9 (greulich near/2 pyle) (6)
- #10 (tanner near/2 whitehouse) (12)
- #11 demirjian (11)
- #12 haavikko (1)

#13 kullman (17)

#14 nortje (9)

#15 liversidge (9)

#16 kvaal (5)

#17 #1 or #2 or #3 or #4 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or

#16 in Trials (197)

Database: PubMed

Search date: 2016-03-14

Search (((publisher [sb]) OR pubstatusaheadofprint)) AND (((age determinat*[Title/Abstract]) OR age estimat*[Title/Abstract]) OR age assess*[Title/Abstract])	46
---	----

Database: Google Scholar

Search date: 2016-03-23

"age estimation" OR "estimation of age" OR "estimating age" OR "age determination" OR "determination of age" OR "determining age" OR "age assessment" OR "assessing age" OR "assessment of age"	Leste første 100 treff.
---	-------------------------

Database: Clinicaltrials.gov

Dato for søk: 2016-03-15

"age estimation" OR "estimation of age" OR "estimating age" OR "age determination" OR "determination of age" OR "determining age" OR "age assessment" OR "assessing age" OR "assessment of age"	16
greulich OR pyle OR demirjian OR haavikko OR kullman OR nortje OR liversidge OR kvaal	14

Database: WHO - International Clinical Trials Registry Platform (ICTRP)

Dato for søk: 2016-03-15

age estimation OR estimation of age OR estimating age OR age determination OR determination of age OR determining age OR age assessment OR assessing age OR assessment of age	20
greulich OR pyle OR demirjian OR haavikko OR kullman OR nortje OR liversidge OR kvaal	2

Appendix 2. Description of included studies with quality assessment

Dedouit F, Auriol J, Rousseau H, Rougé D, Crubézy E, Telmon N. Age assessment by magnetic resonance imaging of the knee: a preliminary study. <i>Forensic Sci Int.</i> 2012;217(1-3):232.e1-7.		
Population: Contry, ethnicity, place and year	We retrospectively reviewed clinical knee MR scans of living subjects of documented age, performed in the radiology departments of our institution (Toulouse, France)	
Age and sex, sample	290 patients aged 10-30 year (138 males, 152 females)	
Design of the study	Retrospective	
Index test	Five stages MR	
Aim of the study	"The aims of our study were firstly to develop an original MRI staging system of growth plate maturation and epiphyseal fusion of the knee for estimation of skeletal bone age in a population ranging from 10 to 30 years old, and secondly to evaluate the reliability and validity of the system."	
QUADAS-2 assessment		
Patient selection method:		
	Rating	Comment
- Consecutive or random sample of patients?	Yes	"We selected randomly and retrospectively 290 patients aged 10–30 years old"
- Avoid inappropriate exclusions?	Yes	
DOMAIN 1: Patient selection	Low risk	
DOMAIN 1: Extra questions on age cohorts and age range	Unclear risk (sample size too small in each age group)	
- Index test interpreted without knowledge of CA?	Yes	"Images were anonymised by deleting the names and documented ages of the patients."
DOMAIN 2: Index test interpretation	Low risk	
- CA interpreted without knowledge of SA?	Yes	We assume age was recorded in the medical records when patients went to the hospital prior to this study
DOMAIN 3: Reference standard	Low risk	
- All patients included in analysis?	Yes	
DOMAIN 4: Patient flow and timing bias	Low risk	

Ekizoglu O, Hocaoglu E, Can IO, Inci E, Aksoy S, Bilgili MG. Magnetic resonance imaging of distal tibia and calcaneus for forensic age estimation in living individuals. Int J Leg Med. 2015;129(4),825-831.		
Population: Contry, ethnicity, place and year	The present study was conducted at Bakirkoy Dr. Sadi Konuk Teaching and Research Hospital (Turkey) from January to November 2013	
Age and sex, sample	167 MR images aged 8-25 years (97 maless, 70 femaless)	
Design of the study	Retrospective	
Index test	Saint-Martin three stages system MR	
Aim of the study	"The aim of this study was to enlarge the database on MR analysis of ossification of the distal tibial and calcaneal epiphysis developed by Saint-Martin et al."	
QUADAS-2 assessment		
Patient selection method:		
	Rating	Comment
- Consecutive or random sample of patients?	Yes	
- Avoid inappropriate exclusions?	Yes	
DOMAIN 1: Patient selection	Low risk	
DOMAIN 1: Extra quesitons on age cohortrs and age range	Unclear risk (sample size too small in each age group)	
- Index test interpreted without knowledge of CA?	Unclear	
DOMAIN 2: Index test interpretation	Unclear risk	
- CA interpreted without knowledge of SA?	Yes	"In this retrospective study, we obtained sex and age information and all medical documents of 198 patients from the data-processing center of the hospital."
DOMAIN 3: Reference standard	Low risk	
- All patients included in analysis?	Yes	
DOMAIN 4: Patient flow and timing bias	Low risk	

Ekizoglu O, Hocaoglu E, Inci E, Can IO, Aksoy S, Kazimoglu C. Forensic age estimation via 3-T magnetic resonance imaging of ossification of the proximal tibial and distal femoral epiphyses: Use of a T2-weighted fast spin-echo technique. Forensic Sci Int. 2016;260:102.e1-7.		
Population: Contry, ethnicity, place and year	"This study was conducted at the Bakirkoy Dr. Sadi Konuk Teaching and Research Hospital (Turkey) from June 2012 to June 2014"	
Age and sex, sample	503 patients aged 10-30 years (305 maless, 198 femaless)	
Design of the study	Retrospective	
Index test	Five stages MR	
Aim of the study	"The purpose of this study is to evaluate the reliability and validity of the five-stage age estimation method defined by Dedouit et al. for distal femoral and tibial epiphyses in Turkish population and to provide the enlargement of database."	
QUADAS-2 assessment		
Patient selection method:		
	Rating	Comment
- Consecutive or random sample of patients?	Yes	
- Avoid inappropriate exclusions?	Yes	
DOMAIN 1: Patient selection	Low risk	
DOMAIN 1: Extra quesitons on age cohortrs and age range	High risk (the numbers in each age group distributed unevenly, some groups had more than 20 subjects while some groups had only 3-5 subjects)	
- Index test interpreted without knowledge of CA?	Yes	
DOMAIN 2: Index test interpretation	Low risk	
- CA interpreted without knowledge of SA?	Unclear	
DOMAIN 3: Reference standard	Unclear risk ("Data on age of the patients were obtained from the hospital database, which included only birth years.")	
- All patients included in analysis?	Yes	
DOMAIN 4: Patient flow and timing bias	Low risk	

Krämer JA, Schmidt S, Jürgens KU, Lentschig M, Schmeling A, Vieth V. Forensic age estimation in living individuals using 3.0 T MRI of the distal femur. Int J Legal Med. 2014 May;128(3):509-14.		
Population: Contry, ethnicity, place and year	"TheMRI scans of 304 individuals in the age group from 10 to 30 years were acquired between 2010 and 2012 at the Center ofModernDiagnostics (ZEMODI) in Bremen (Germany)"	
Age and sex, sample	290 individuals aged 10-30 years (166 maless, 124 femaless)	
Design of the study	Retrospective	
Index test	Five stages MR	
Aim of the study	"Forensic age estimation in living individuals using 3.0T MRI of the distal femur"	
QUADAS-2 assessment		
Patient selection method:		
	Rating	Comment
- Consecutive or random sample of patients?	Yes	
- Avoid inappropriate exclusions?	Yes	
DOMAIN 1: Patient selection	Low risk	
DOMAIN 1: Extra quesitons on age cohortrs and age range	Unclear risk (sample size too small in each age group)	
- Index test interpreted without knowledge of CA?	Unclear	
DOMAIN 2: Index test interpretation	Unclear risk	
- CA interpreted without knowledge of SA?	Yes	We assume age was recorded in the medical records when patients went to the hospital prior to this study
DOMAIN 3: Reference standard	Low risk	
- All patients included in analysis?	Yes	All participants were included in the study, but the authors only presented data in stage 2 and 3 in the published paper
DOMAIN 4: Patient flow and timing bias	Low risk	

Appendix 3. Separate list

In addition to the 4 included studies above, we identified 4 studies with relevant study design but no available data. These studies were listed below.

Title	Country	Year	Age	Sex	Size	Material	Methods	Reason
Fan F, Zhang K, Peng Z, Cui JH, Hu N, Deng ZH. Forensic age estimation of living persons from the knee: Comparison of MRI with radiographs. <i>Forensic Sci Int.</i> 2016;268:145-150.	China	2016	11-?	Both	322	Knee	X-ray & MRI	Sample size in each stage is not presented
Jopp E, Schröder I, Maas R, Adam G, Püschel K. Proximales Tibiaepiphyse im Magnetresonanztomogramm: Neue Möglichkeit zur Altersbestimmung bei Lebenden? <i>Rechtsmedizin.</i> 2010;20(6):464-468.	Germany	2010	15-19	Both	41	Knee	MRI	Relevant data is not presented
Saint-Martin P, Rérolle C, Dedouit F, Bouilleau L, Rousseau H, Rougé D, Telmon N. Age estimation by magnetic resonance imaging of the distal tibial epiphysis and the calcaneum. <i>Int J Legal Med.</i> 2013;127(5):1023-30.	France	2013	8-25	Both	180	Ankle	MRI	Relevant data is not presented
Saint-Martin P, Rérolle C, Dedouit F, Rousseau H, Rougé D, Telmon N. Evaluation of an automatic method for forensic age estimation by magnetic resonance imaging of the distal tibial epiphysis--a preliminary study focusing on the 18-year threshold. <i>Int J Legal Med.</i> 2014;128(4):675-83.	France	2014	8-25	Both	160	Ankle	MRI	Relevant data are not presented (data may be duplicated as Saint-Martin 2013)

Appendix 4. Excluded studies

Author	Reason
Aly SM, Shrestha B, Hong DJ, Omran A, Wang W. Identification of age and sex based on knee radiography. <i>Forensic Sci Int.</i> 2016;267:231.e1-7.	Not CT or MR
Galić I, Mihanović F, Giuliadori A, Conforti F, Cingolani M, Cameriere R. Accuracy of scoring of the epiphyses at the knee joint (SKJ) for assessing legal adult age of 18 years. <i>Int J Legal Med.</i> 2016;130(4):1129-42.	Not CT or MR
Banerjee KK, Agarwal BB. Estimation of age from epiphyseal union at the wrist and ankle joints in the capital city of India. <i>Forensic Sci Int.</i> 1998;98(1-2):31-9.	Not CT or MR
Bokariya P, Chowdhary DS, Kothari R. Estimation of age of individual by radiographic examinations of bones at knee joint in western rajasthan population. <i>J Foren Med Toxi.</i> 2009;26(2):1-3.	Not CT or MR
Cameriere R, Cingolani M, Giuliadori A, De Luca S, Ferrante L. Radiographic analysis of epiphyseal fusion at knee joint to assess likelihood of having attained 18 years of age. <i>Int J Legal Med.</i> 2012;126(6):889-99.	Not CT or MR
Crowder C, Austin D. Age ranges of epiphyseal fusion in the distal tibia and fibula of contemporary maless and femaless. <i>J Forensic Sci.</i> 2005;50(5):1001-7.	Not CT or MR
Davies C, Hackman L, Black S. The persistence of epiphyseal scars in the adult tibia. <i>Int J Legal Med.</i> 2014;128(2):335-43.	No age estimation
Dogaroiu C, Avramoiu M. Correlation between chronological age and the stage of union of the distal femur and proximal tibia epiphyses in a Romanian sample population. <i>Rom J Leg Med.</i> 2015;23(3):171-176.	Not CT or MR
Faisant M, Rerolle C, Faber C, Dedouit F, Telmon N, Saint-Martin P. Is the persistence of an epiphyseal scar of the knee a reliable marker of biological age? <i>Int J Legal Med.</i> 2015;129(3):603-8.	Not CT or MR
Garnov N, Gründer W, Thörmer G, Trampel R, Turner R, Kahn T, Busse H. In vivo MRI analysis of depth-dependent ultrastructure in human knee cartilage at 7 T. <i>NMR Biomed.</i> 2013;26(11):1412-9.	Not age estimation
Hackman L, Black S. Age estimation from radiographic images of the knee. <i>J Forensic Sci.</i> 2013;58(3):732-7.	Not CT or MR

Hackman L, Davies CM, Black S. Age estimation using foot radiographs from a modern Scottish population. <i>J Forensic Sci.</i> 2013;58 Suppl 1:S146-50.	Not CT or MR
Kanchan T, Krishan K. Radiographic assessment of age from epiphyseal fusion at knee joint. <i>Int J Legal Med.</i> 2013;127(4):839-41.	Comments, not an empirical study
Kong CK, Tse PW, Lee WY. Bone age and linear skeletal growth of children with cerebral palsy. <i>Dev Med Child Neurol.</i> 1999;41(11):758-65.	Not age estimation, sample size less than 50
Mosle HG. Methods for the determination of the age of the femur. <i>Deutsche Zeitschrift fuer die gesamte gerichtliche medizin.</i> 1966;58(4):205-11.	Not age estimation
O'Connor JE, Bogue C, Spence LD, Last J. A method to establish the relationship between chronological age and stage of union from radiographic assessment of epiphyseal fusion at the knee: an Irish population study. <i>J Anat.</i> 2008 ;212(2):198-209.	Not CT or MR
O'Connor JE, Coyle J, Bogue C, Spence LD, Last J. Age prediction formulae from radiographic assessment of skeletal maturation at the knee in an Irish population. <i>Forensic Sci Int.</i> 2014;234:188.e1-8.	Duplicated data & not CT or MRI
O'Connor JE, Coyle J, Spence LD, Last J. Epiphyseal maturity indicators at the knee and their relationship to chronological age: results of an Irish population study. <i>Clin Anat.</i> 2013;26(6):755-67.	Not CT or MR
Saint-Martin P, Rérolle C, Pucheux J, Dedouit F, Telmon N. Contribution of distal femur MRI to the determination of the 18-year limit in forensic age estimation. <i>Int J Legal Med.</i> 2015;129(3):619-20.	Comments, not an empirical study
Schmidt S, Vieth V, Timme M, Dvorak J, Schmeling A. Examination of ossification of the distal radial epiphysis using magnetic resonance imaging. New insights for age estimation in young footballers in FIFA tournaments. <i>Sci Justice.</i> 2015;55(2):139-44.	Not knee or ankle
Shah J, Howale D, Patel D, Patel V. Age determination from epiphyseal union at knee joint. <i>J Forensic Toxicol.</i> 2011;28(2):19-23.	Not CT or MR
Tajima T. A method of assessing skeletal maturity of the knee joint. <i>Nihon Seikeigeka Gakkai Zasshi.</i> 1964;38:791-804.	Not CT or MR
Xue XJ, Tian XM, Zhang JZ, Min JX, Liu XD. Estimation of age and radiographic changes in Chinese girls at the age of 11-16 years with the characteristic of X-ray image of articulations of bone. <i>Chin J Forensic Med.</i> 2006;21(1):22-23.	Not CT or MR

Appendix 5. Abbreviations

CA: chronological age

CI: confidence interval

CT: computed tomography

GRADE: Grading of Recommendations Assessment, Development and Evaluation

MRI: magnetic resonance imaging

QUADAS: Quality Assessment of Diagnostic Accuracy Studies

SD: standard deviation

Appendix 6. Project protocol

Note: This protocol is published on December 2016, which is a combination for several projects, including age estimation with hand, third molar teeth, clavicle and knee and ankle.

We will conduct systematic reviews of the precision and accuracy of methods for age estimation of young people (10 to 25 years of age).

Literature search

Research librarian, Gyri Hval Straumann (GHS), will plan the searches in collaboration with the project group. Another research librarian will peer review the search strategy. GHS will execute the electronic searches for literature in the following databases:

- MEDLINE
- Embase
- Cochrane Central Register of Controlled Trials (CENTRAL)
- Google scholar

In addition, we will search for ongoing trials in Clinicaltrials.gov and WHO International Clinical Trials Registry Platform, and check references and citations of relevant articles in Web of Science.

Inclusion criteria

Study design:	We will include studies that compare an age indicator of one of the index tests specified below with known chronological age.
Population:	Living persons between the age of 10 and 25 years old
Index test:	Age estimation tests including: <ul style="list-style-type: none">• X-ray images of the hand-wrist• dental X-ray images• CT or MR of the clavicle, knee or ankle
Reference test:	Known chronological age
Outcome:	For x-ray images of the hand: The difference between the estimated age and chronological age. For x-ray images of the teeth or CT or MR of the clavicle, knee or ankle: The difference between the estimated age and chronological age, or the distribution of chronological age given a specific age indicator (typically a pre-defined development stage).
Language:	No restrictions in the search

Exclusion criteria

We will exclude studies that have fewer than 50 persons that qualify for our analysis.

Article selection

At least two reviewers will independently screen each of the identified references. We will conduct one combined literature search. Because we expect many studies relevant for inclusion, we plan to separate the potentially relevant references into three groups for full-text screening where we will handle each index test separately, resulting in three systematic reviews. At least two authors will independently screen each of the full-text references considered potentially relevant.

Assessment of methodological quality and data extraction

At least two authors will independently use QUADAS-2 to assess methodological quality of each of the included studies. We will resolve any disagreement by discussion or by consulting a third reviewer. One person will collect data about population (age, ethnicity and sex), where and when the studies were conducted, the reference and index tests used, and the results. Another person will check the extraction.

Statistical methods

Examination of effect size

In the current project, we would like to estimate the mean chronological age and its variation given a certain hand-wrist or teeth radiograph results, and clavicles, knee and ankles MRI and CT results. Therefore, we will present the mean chronological age and its 95% confidence interval separately in both gender (male/female) and in all the skeleton age groups or age category.

We will use random effects models to pool summary data at trial level. Potential heterogeneity of country and publication date will be explored in Meta-regression. Publication bias will be assessed with Funnel plots and Egger's regression model (1).

All results will be checked for statistical heterogeneity using the Cochrane Q statistic ($P < 0.10$) and I^2 methodology ($I^2 > 50\%$). Statistical significance is set at a P value < 0.05 . The meta-analysis will be performed using R software (R Core Team, Vienna, Austria; version 3.3.2; using packages “glm”, “lme4” and “metafor”).

References

1. Egger, M., et al., Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 1997. 315(7109): p. 629-34.

Published by the Norwegian Institute of Public Health

August 2018

P.O.B 4404 Nydalen

NO-0403 Oslo

Telefon: + 47-21 07 70 00

The report can be downloaded as pdf
at www.fhi.no/en/publ/