Article:
Shalof, H., Dimitri, P., Shuweihdi, F. et al. (1 more author) (2021) "Which skeletal imaging modality is best for assessing bone health in children and young adults compared to DXA? A systematic review and meta-analysis". Bone, 150. 116013. ISSN 8756-3282

https://doi.org/10.1016/j.bone.2021.116013

‘’ Which skeletal imaging modality is best for assessing bone health in children and young adults compared to DXA? A systematic review and meta-analysis’’

Authors and affiliations

Heba Shalof[1,2], Paul Dimitri[1,3], Farag Shuweihdi[4], Amaka C Offiah[1,5]

1 Academic Unit of Child Health, Department of Oncology and Metabolism, University of Sheffield, Damer Street Building, Western Bank, Sheffield S10 2TH, United Kingdom.

2 Faculty of medicine, Omar Al-Mukhtar University, Bayda, Libya.

3 Department of Pediatric Endocrinology, Sheffield Children’s NHS Foundation Trust, Western Bank, Sheffield, United Kingdom.

4 Leeds Institute of Health Sciences, School of medicine, University of Leeds, Leeds, United Kingdom.

5 Radiology Department, Sheffield Children’s NHS Foundation Trust, Western Bank,

Sheffield, United Kingdom.

Corresponding author:

Mrs. Heba Shalof

E-mail: hsmshalof1@sheffield.ac.uk.

University of Sheffield

Academic Unit of Child Health

Damer Street Building

Western Bank

Sheffield

S10 2TH

ORCID ID:

HS: 0000-0003-4659-1363

PD: 0000-0001-7625-6713

FS: 0000-0003-1199-2992

ACO: 0000-0001-8991-5036

Declarations of Interest

None

Abstract:

Background Skeletal imaging techniques have become clinically valuable methods for measuring and assessing bone mineral density in children and young people. Dual-energy X-ray absorptiometry (DXA) is the current reference standard for evaluating bone density, as recommended by the International Society for Clinical Densitometry (ISCD). Various bone imaging modalities, such as quantitative ultrasound (QUS), peripheral quantitative computed tomography (pQCT), high-resolution peripheral quantitative computed tomography (HR-pQCT), magnetic resonance imaging (MRI), and digital X-ray radiogrammetry (DXR) have been developed to further quantify bone health in children and adults. The purpose of this review, with meta-analysis, was to systematically research the literature to compare the various imaging methods and identify the best modality for assessing bone status in healthy papulations and children and young people with chronic disease (up to 18years).

Methods A systematic computerized search of Medline, PubMed, and Web of Science databases was conducted to identify English-only studies published between 1st January 1990 and 1st December 2019**.** In this review, clinical studies comparing imaging modalities with DXA were chosen according to the inclusion criteria. The risk of bias and quality of articles was assessed using the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS-2). The meta-analysis to estimate the overall correlation was performed using a Fisher Z transformation of the correlation coefficient. Additionally, the diagnostic accuracy measures of different imaging methods compared with DXA were calculated.

Results The initial search strategy identified 13,412 papers, 29 of which matched the inclusion and exclusion criteria. Of these, twenty-two papers were included in the meta-analysis. DXA was compared to QUS in 17 papers, to DXR in 7 and to pQCT in 4 papers. A single paper compared DXA, DXR, and pQCT. The meta-analysis demonstrated that the strongest correlation was between DXR and DXA, with a coefficient of 0.71 [95%CI: 0.43; 1.00, p-value<.001]), while the correlation coefficients between QUS and DXA, and pQCT and DXA were 0.57 [95%CI: 0.25; 0.90, p-value<.001] and 0.57 [95%CI: 0.46; 0.67, p-value<.001], respectively. The overall sensitivity and specificity were statistically significant 0.71 and 0.80, respectively.

Conclusion No current imaging modality provides a full evaluation of bone health in children and young adults, with each method having some limitations. Compared to QUS and pQCT, DXR achieved the strongest positive relationship with DXA. DXR should be further evaluated as a reliable method for assessing bone health and as a predictor of fractures in children and young people.

Highlights:

- No current imaging modality provides a full evaluation of bone health in children and young adults, with each method having some limitations.
- To date, dual-energy X-ray absorptiometry (DXA) is the reference standard for evaluating BMD in children and adults, as recommended by the International Society for Clinical Densitometry (ISCD).
- From meta-analysis, the strongest relationship was between digital X-ray radiogrammetry (DXR) and DXA, with a pooled estimate of 0.71.

Keywords:

Bone health; Fracture risk; Bone mineral density; Systematic review; Bone densitometry; Children.

Abbreviations

aBMD	areal Bone Mineral Density
Ad-SoS	Amplitude dependent speed of sound
BHI	Bone Health Index
BMAD	Bone Mineral Apparent Density
BMD	Bone Mineral Density
BMD_{FN}	Bone Mineral Density Femoral neck
BMD_{LS}	Bone Mineral Density Lumbar Spine
BMD_{TB}	Bone Mineral Density Total Body
BTT	Bone Transmission Time
BUA	Broadband Ultrasound Attenuation
CoTh	Cortical Thickness
CSA	Cross-Sectional Area
CSMI	Cross-Sectional Moment of Inertia
DXA	Dual-energy X-ray Absorptiometry
DXR	Digital X-ray Radiogrammetry
FN	False negative
FP	False positive
HR-pQCT	High-Resolution peripheral Quantitative Computed Tomography
ISCD	International Society for Clinical Densitometry
MRI	Magnetic Resonance Imaging
PA-BMD	Posteroanterior bone mineral density
pQCT	Peripheral Quantitative Computed Tomography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
QCT	Quantitative Computed Tomography
QUS	Quantitative bone Ultrasound
QUADAS-2	Quality Assessment Tool for Diagnostic Accuracy Studies
RE model	Random-Effects model

ROI	Region of Interest
SDS	Standard deviation score
SOS	Speed of Sound
TN	True Negative
TP	True Positive
vBMD	volumetric Bone Mineral Density
WA-BMD	Width-adjusted bone mineral density

Introduction

The ability to assess bone strength and measure bone mineral density in children has greatly evolved during the past four decades. The aim of these assessments is to identify children with a high risk of low-trauma fractures due to bone fragility or chronic illness, manage their treatment decisions, and monitor their therapy responses [1].

Bone health is a general concept that describes the healthy and nutritional status of bone. It represents the impact of various conditions and factors that contribute to bone weakness or strength, i.e., an increase or decrease in bone fragility. The cornerstone for evaluating bone health is the bone's ability to resist fracture, known as bone strength. Bone strength is influenced by its material properties, such as bone mineral density and bone matrix composition (organic and non-organic components). Other factors should be considered when assessing bone strength, such as bone microarchitecture and bone geometry. The Consensus Development Panel on Osteoporosis in the National Institute of Health (NIH) recognized the role of bone microarchitecture as an essential factor in the assessment of bone health [2].

Multiple imaging modalities have been developed to measure bone mineral density (BMD) and evaluate bone health. To date, dual-energy X-ray absorptiometry (DXA) is the recognized reference standard for assessing BMD in children and adults, as recommended by the International Society for Clinical Densitometry (ISCD) [3, 4]. It has become one of the most clinically valuable tools in the assessment of pediatric bone health, due to its wide availability and being inexpensive, easy to use and low radiation dose (1–6 μSv) [5, 6].

However, there are limitations to DXA. For example, it is less reliable in children below the age of 4 years, or who are small for their age, because DXA measurements are affected by bone size [7-9]. DXA is unable to measure bone depth (true volumetric bone mineral density) and its results depend on a two-dimensional projection of a three-dimensional structure. Therefore, the size dependence of aBMD impedes utilizing DXA in those children suffering from abnormal growth patterns, delayed sexual maturation and chronic disorders [10, 11]. Many studies conclude that areal bone mineral density (aBMD) is unable to predict future fracture risk in children [12-14]. Furthermore, DXA cannot differentiate between cortical and trabecular bone [15]. Hence, it cannot provide detailed information on bone microarchitecture and skeletal integrity and strength.

To address the need to evaluate bone microarchitecture, other bone imaging modalities, including quantitative ultrasound (QUS), peripheral quantitative computed tomography (pQCT), high-resolution peripheral quantitative computed tomography (HR-pQCT), magnetic resonance imaging (MRI), and digital x-ray radiogrammetry (DXR), have been developed to quantify bone morphometry and strength in children. Each imaging modality has specific limitations in the assessment of bone strength and measurement of BMD, with challenges for accurately predicting fracture risk. These modalities have been applied to research, but currently, there is no reliable clinical imaging modality for predicting fracture risk and assessing bone strength in children.

QUS may be advantageous in evaluating and assessing bone health in young children due to the technique being safe (radiation-free), easy and portable [16-18]. Despite the advantages of QUS for measuring and evaluating bone health in children, infants, and preterm infants, utilization of this method is still controversial. Some researchers have suggested that bone size may affect the accuracy of measurement of QUS, especially in growing children and the site of measurement may not reflect the density in other parts of the skeleton [19]. In addition, a lack of pediatric reference data may lead to some difficulties when comparing the data between different studies.

QCT has become widely used in bone research in adults because it can evaluate cortical and trabecular bone separately, and compute true volumetric bone mineral density (vBMD in g/cm^3) [20]. However, QCT has some limitations, especially when used in children. The radiation dose of QCT is relatively high (approximately 90 μSv) [21], and limited pediatric reference data exist [22]. pQCT was thus developed for use in the evaluation of bone health to overcome QCT limitations, especially among children. In principle it is similar to QCT, but the equipment is smaller, more mobile and the radiation dose (10 μSv) is lower than that from QCT. However, pQCT can only assess the appendicular skeleton, whereas QCT can assess both the axial and the appendicular skeleton [23].

Subsequently HR-pQCT was developed as a non-invasive and low radiation approach for measuring volumetric bone mineral density, assessing cortical and trabecular compartments separately, evaluating micro architectural morphology and using finite element analysis to determine bone strength [24]. HRpQCT is used to image the peripheral sites with an acquisition resolution of 82 μm [a new generation HR-pQCT scanner (Scanco Xtreme CT II) has been developed with a resolution of 58 μm], while pQCT resolution is approximately 400 μm. There are some limitations in using pQCT and HR-pQCT, especially in children. For example, movement artifact which has a considerable effect on bone microarchitecture [25]. The reason for this artifact is a product of long scan time (7-15 min for pQCT, 3 min for HR-pQCT) [26]. Currently, only relatively few HR-pQCT scanners are present in hospitals and the lack of standardized pediatric references data for the micro-architectural and volumetric BMD parameters (especially for children below the age of eight years) restrict the widespread clinical use of this technique.

DXR evaluates bone mineral density from hand radiographs [27]. One automated method uses the computer software BoneXpert, developed for measuring and calculating bone age and bone mass. This software measures the cortical thickness, length, and width of the three middle metacarpals from hand radiographs, with results expressed as the bone health index (BHI) [28]. DXR is considered a safe tool for evaluating bone health in pediatric populations

because its radiation dose is low (about <1 μSv) and is applied to an extremity. Although DXR is an inexpensive, widely available method which can easily be used in children, more work is needed to ascertain whether DXR is able to reliably evaluate bone status and to enable it to be used as a standard tool to predict fractures in children.

Our aim was to systematically review the literature to compare the various currently available imaging modalities with DXA (being the reference standard) and to identify the most reliable for assessing and evaluating bone health in children and young adults.

Materials and methods

Search strategy and protocol

A systematic review was undertaken according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis checklist (PRISMA) [29]. A systematic computerized search of Medline, PubMed, and Web of Science databases was conducted to identify clinical studies that assessed and compared diagnostic accuracy and/or correlation coefficients of bone health parameters of two or more imaging modalities, with DXA being one of them. All database searches were limited to children and adolescents (healthy and non-healthy) up to 18 years old. The search was restricted to full-text articles, human studies published in the English language between 1st January 1990 and 1st December 2019. The full search strategy can be seen in supplementary materials **(Suppl 1).**

Study selection

Study inclusion and exclusion criteria were developed based on population, intervention, comparator, and outcomes (PICO) criteria.

Inclusion Criteria

- The population was healthy and non-healthy, male and female children and adolescents up to18 years old.
- Clinical studies that assessed and compared diagnostic accuracy and/or correlation coefficient (r) of bone health parameters of two or more imaging modalities, DXA being one of them.
- Only full text publications in English.

Exclusion criteria

- Studies including adults and children where data for children was not separately extractable.
- Studies that did not include DXA as an imaging modality.
- Full text not in English.
- The studies were review articles, case reports, clinical or conference reports.

The study selection was performed in two stages; firstly, one reviewer <<anonymized>> scanned the titles and abstracts of all identified electronic database citations after removing duplicate papers. Articles that did not fulfil the selection criteria were removed. References were arranged, and duplicates were eliminated using EndNote® X9. In the second stage, <<anonymized>> identified potentially relevant papers by reading the full texts. The final list of articles was critically assessed by two reviewers <<anonymized>>. Any disagreement was resolved by consensus.

Data items and data extraction process

The required information was collected into an Excel spreadsheet (**Suppl 2**). The following data were extracted from each paper: First author; year of publication; place of study; study design; aim of study; sample size; mean age (years); main clinical presentation; imaging modalities; parameters and region of interest (ROI) measured; duration of investigation; time interval between investigations; diagnostic accuracy of each modality (if stated); correlation between modalities; radiation dose for each modality. Two reviewers <<anonymized>> reviewed the data extracted from the included papers.

Quality assessment process

The quality of the retrieved papers was assessed independently by three reviewers <<anonymized>> using the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS-2) [30]. This tool consists of four key domains: patient selection; index test; reference standard; flow and timing. The reviewers individually recorded the risk of bias and applicability concerns as "low risk," "high risk," or "unclear risk". The final quality assessment sheet available (**Suppl 3**). Disagreements between reviewers were resolved by consensus. In this review, the maximum acceptable time interval between the standard test (DXA) and the index test (DXR, QUS or pQCT) was set as three months.

Data synthesis and the meta-analysis

The meta-analysis to estimate overall correlation was performed using a Fisher Z-transformation of the correlation coefficient (r) [31]. Because r covers the whole range of relationship strengths, from no relationship (zero) to a perfect relationship (1, or -1), it is telling us exactly how large the relationship really is between the variable we have studied. Cohen provided rules of thumb for interoperating these effect sizes, suggesting that an r of 0.1 represents a 'small' effect size, 0.3 represents a 'medium' effect size and 0.5 represents a ' large' effect size [32]. The pooled effect sizes for correlations were estimated via random-effects (RE) model [33]. Forest plots were used to present the effects of studies and the pooled effect estimate [34]. The heterogeneity between studies was examined using two statistics: The Q and I^2 statistics. For detecting outliers and influencers, sensitivity analysis of the meta-analyses was performed using different visual approaches [35]. To investigate publication and other bias in this analysis, funnel plots based on standard errors was used as a visual tool. Meta-analysis for estimating both sensitivity and specificity were based on extracting the values of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) from these three studies. The statistical package *R* version 4.0.3 was used to perform the meta-analysis.

Results

Literature search results and study selection

The initial search strategy identified 13412 papers from PubMed (n=7123), Medline (n=1943) and Web of Science (n=4337). Of these, 2881 were duplicates and excluded by EndNote® X9. A thorough evaluation of titles and abstracts of 10531 papers was performed by one reviewer <<anonymized>> and 10438 papers were eliminated for the following reasons: irrelevant papers (n =8363); review articles (n =599); clinical or conference report abstracts (n =197); guidelines (n =74); follow up case study (n =382); adult population (n =701); animal study (n =42) or full text not in English (n = 80). The remaining 93 papers were screened with a full text review. Of these, 64 were excluded from this review: not comparison studies (n=13); the study's sample was adult and children and data for children was not separately extractable (n=18); conference reports (n=4); studies that were comparison studies between two imaging modalities where DXA was not one of them (n=11); clinical studies that focused on arthritis and not on bone density measurement and fracture risk (n=5); papers that did not present diagnostic accuracy (sensitivity and specificity) or correlation index between imaging methods or did not present enough data to determine correlation (n=9); and full text not in English (n=4). Twenty-nine papers were relevant and met the inclusion and exclusion criteria.

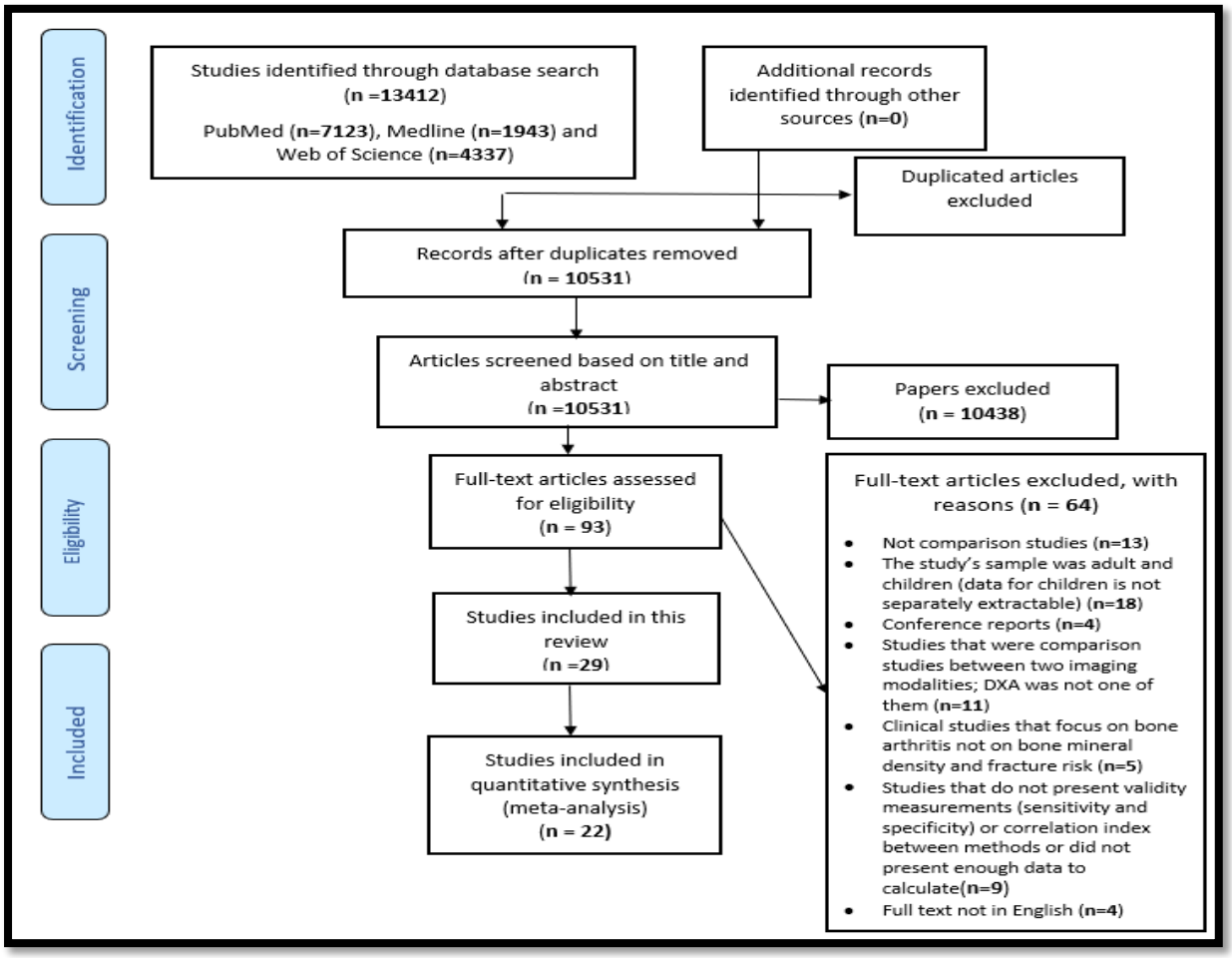

Figure 1: PRISMA flowchart summarizing the search

Quality assessment of studies

Most papers were low risk for methodology; the percentage (%) of papers that showed a low risk of bias in each of the four domains (patient selection, index test, reference standard, and flow and timing) were 76%, 52%, 48%, and 35%, respectively while the percentage designated as high risk were 10%, 6%, 6%, and 13%, respectively. The remainder were scored as "unclear" because the authors did not offer enough details about index and reference test results and how they interpreted them, as well as poor reporting about the interval between two index and reference methods.

Regarding applicability, all papers were of low concern except one [36], which was scored as being of high concern in patient selection and standard test domains due to 38 participants having only DXA scans, 335 healthy children having only quantitative ultrasound scans and 8 participants having both scans without the exact reasons for this being detailed.

Study characteristics

The 29 papers included were published between 2000 and 2019. The total sample size was 4221 (range from 22 [37] to 511 [38]). DXA was compared to QUS in 17 papers, to DXR in 7 and to pQCT in 4 papers. A single paper compared DXA, DXR, and pQCT [39]. No paper included HRpQCT.

Analysis of studies that compared Digital X-ray Radiogrammetry (DXR) and dual-energy X-ray absorptiometry (DXA)

The degree of correlation between DXA and DXR/BHI was compared in seven studies [28, 40-45]. There was a moderate to strong degree of correlation between DXA and DXR/BHI in four (**Table 1**). In contrast, one study concluded that the degree of correlation was poor [40], while another study found no correlation [44].

Table 1: The degree of correlation between DXA and DXR*

Authors	Year	Sample Size	Correlation/Conclusion
Mentzal et al	2006	26	DXR-BHI significantly correlated with DXA-BMD_{LS} (r= 0.78)
Van rijn et al	2006	67	There is strong correlation between DXR BHI and DXA BMD_{LS} for acute lymphoblastic leukemia and growth hormone deficiency groups (r= 0.853 M, 0.760 F and 0.760 M, 0.779 F) respectively
Nusman et al	2015	35	Weak correlation between DXA and DXR. The Pearson correlation coefficient for DXA BMD_{LS} Z-score with BHI = 0.247
Neelis et al	2018	46	Spearman's correlation between BHI and BMD_{TB} Z-scores was 0.856 (p <0.001)
Alshamrani et al	2019	465	Weak correlation was found between BHI SDS and DXA BMD_{TB} and BMD_{LS} Z-scores (r=0.11 to r=0.35).
Leijten et al	2019	101	All BHI Z-scores were moderately correlated with DXA BMD_{LS} Z-score (r= 0.564, p < 0.001)

* Mulugeta et al. study, did not address the degree of correlation between DXR and DXA. It only reflects the diagnostic accuracy of these methods.

All the Fisher's transformed correlations used in the forest plots were converted back to correlation to be within the range -1 and +1. There was a positive relationship between DXR and DXA; using pooled estimates this was 0.71 [95%CI: 0.405 1.00, p-value<.001] which is strong and statistically significant. **(Figure 2**).

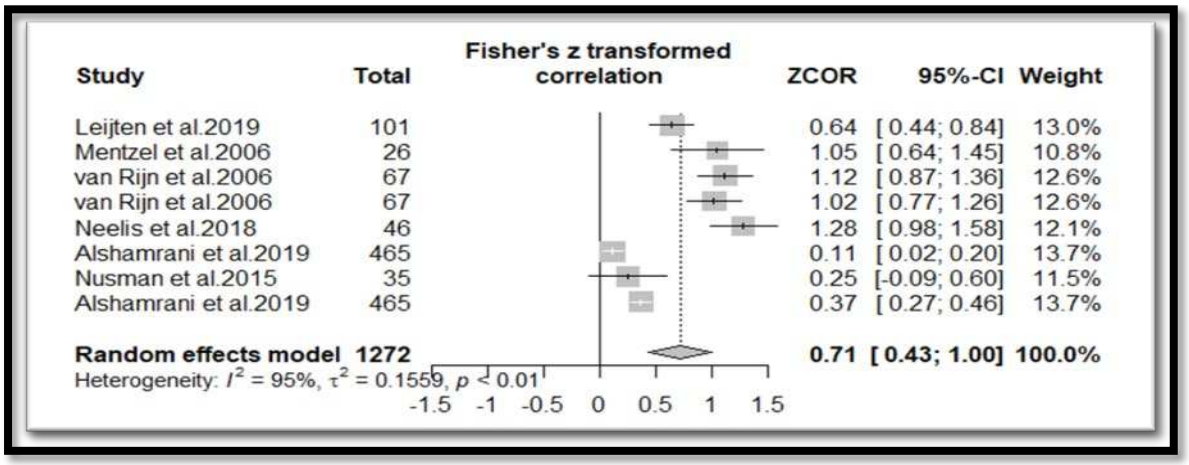

Figure 2: Forest plot of 6 studies for relationship between DXR and DXA using Fisher's transformed correlation

The level of heterogeneity was high (I^2= 95%), and the *Q*-test was statistically significant (p-value<0.001). Regarding the sensitivity influence in terms of outliers, no potential outliers were detected. This means that none of the previous studies had a very high or very small correlation coefficient which might have negatively affected the pooled estimated size of the correlation.

For bias publication, we noted that the largest studies occupied the top of the graph with the smallest standard errors (**Figure 3**). Most of the studies were out of diagonal lines for the 95% confidence limits around the overall correlation.

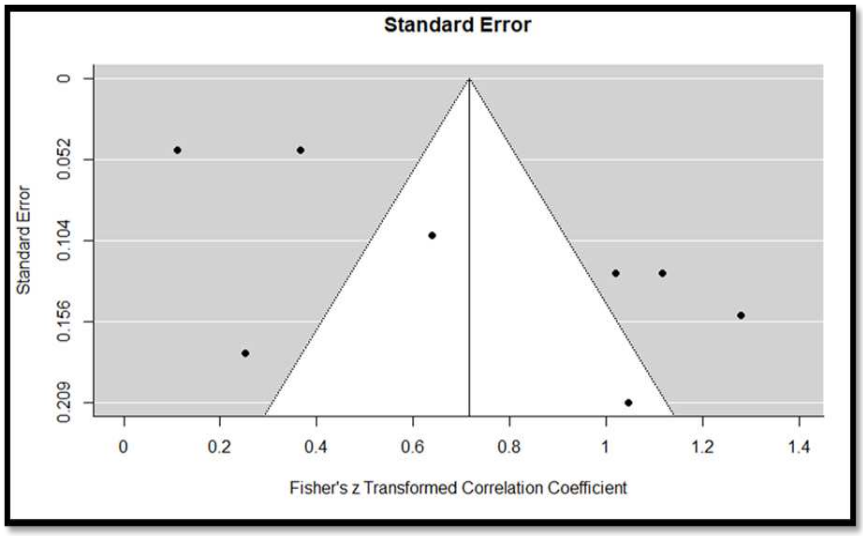

Figure 3: Funnel plot for DXR and DXA

Only three papers calculated the sensitivity and specificity of bone health index (BHI) values comparing them with DXA parameters [28, 41, 44] (**Table 2**). Each of these studies used different DXA parameters (DXA_{LS} Z-score, DXA BMD_{TB}, DXA BMD_{LS}, and DXA BMAD) to compare BHI, and to calculate its sensitivity and specificity.

Table 2: Sensitivity and specificity of BHI compared to DXA.

Authors	Year	Sample Size	Sensitivity	Specificity	DXA parameters
Nusman et al	2015	35	40%	60%	compared to DXA $BMAD_{LS}$
Neelis et al	2018	46	90% 60% 60%	86% 79% 93%	compared to DXA BMD_{TB} compared to DXA BMD_{LS} compared to DXA $BMAD_{LS}$
Leijten et al	2019	101	67%	83%	compared to DXA_{LS} Z-scores.

<u>Analysis of studies that compared Peripheral Quantitative Computed Tomography (pQCT) and dual-energy X-ray absorptiometry (DXA)</u>

Four papers compared pQCT and DXA in the measurement of bone strength and bone density in healthy and unhealthy children and young people [46-49]. Only three calculated the Pearson's correlation coefficient (**Table 3**).

Table 3: The degree of correlation between DXA and pQCT

Authors	Year	Sample Size	Correlation
Kalkwarf et al	2011	444	There was weak to moderate correlation between bone Z-scores for pQCT measures and those from DXA (r=0.11 to 0.60)
Tsampalieros et al [49]	2014	65	pQCT vBMD Z-scores were weakly to moderately correlated with changes in PA-BMD (r=0.36, p<0.01), PA-BMD (r=0.33, p<0.01), and WA-BMD (r=0.49, p<0.001) Z-scores
DiVasta et al	2016	202	There was positive correlation between trabecular and cortical vBMD Z-scores of pQCT and DXA BMD Z-scores (r range 0.57–0.82)

Using pooled estimates, the positive relationship for pQCT and DXA in the previous three studies was 0.57 [95%CI: 0.25; 0.90, p-value <0.001] **(Figure 4)**. The level of heterogeneity was high (I^2= 97%), and the Q-test was statistically significant (p-value <0.0011).

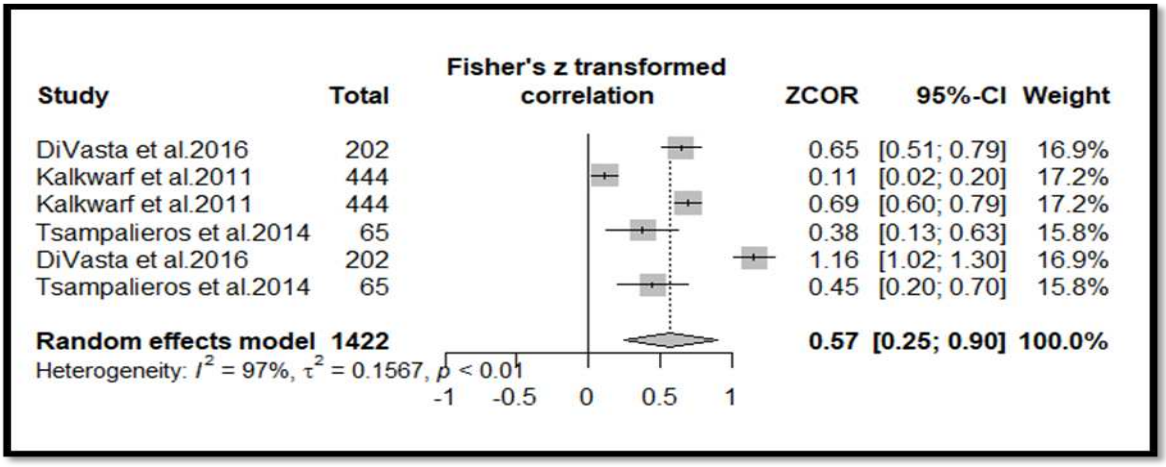

Figure 4: Forest plot of 3 studies for relationship between pQCT and DXA using Fisher's transformed correlation

With respect to the sensitivity analysis of outliers, one study [46] seemed to have slightly low influence, excluding that value did not affect the overall estimate of correlation, and hence it was retained in the analysis. For bias publication, we noted that the largest studies have the smallest standard errors (**Figure 5**). Half of the studies were within the diagonal lines for the 95% confidence limits around the pooling.

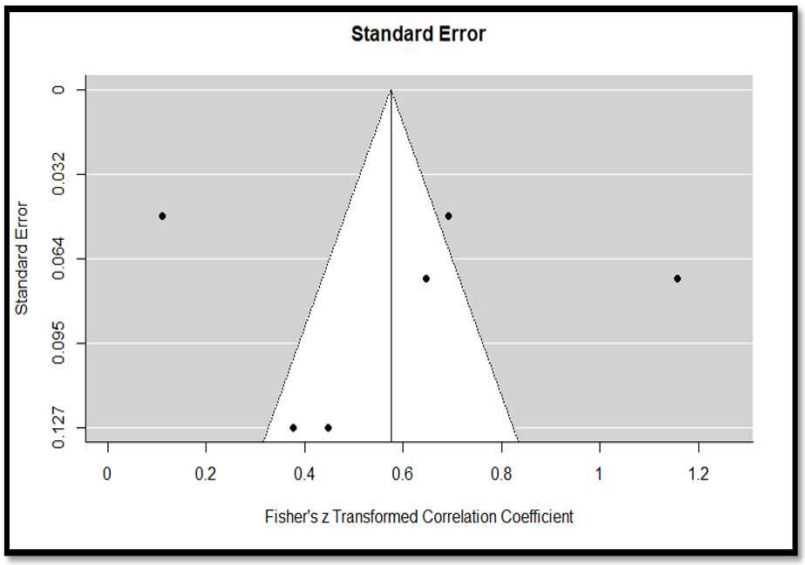

Figure 5: Funnel plot for pQCT and DXA

Analysis of studies that compared Quantitative Ultrasound (QUS) and dual-energy X-ray absorptiometry (DXA)

Of the seventeen studies that compared quantitative ultrasound (QUS) with DXA, 6 demonstrated strong correlation [37, 38, 50-53], 3 showed moderate correlation [54-56], and 6 showed no correlation [57-62] (**Tables 4,5,6**).

Table 4: Studies showing a strong correlation between QUS and DXA.

Authors	Year	Sample Size	Correlation	Comments
Falcini et al	2000	108	BUA with BMD_{LS} was (r =0.83) BUA and BMC_{LS} was (r= 0.81)	BUA was significantly correlated with BMD_{LS} and BMC_{LS}
Njeh et al	2000	22	BMD_{LS} and Ad-SoS was (r = 0.57) BMD_{TB} and Ad-SoS was (r = 0.68)	BMD_{LS} and BMD_{TB} correlated significantly with tibia SOS
Hartman et al	2004	40	BMD_{LS} and SOS at the radius (r = 0.54). BMD_{LS} and SOS at the tibia (r = 0.26)	There was a significant correlation between BMD_{LS} and SOS at the radius and weak correlation with SOS at the tibia
Aceto et al	2014	T0 =30 T1= 30 T2=18 T3=12	AD-SoS Z-score (r=0.86, 0.345, 0.808, −0.368, −0.09). BTT Z-score (r=0.519, 0.530, 0.093, 0.405, 0.754,). Respectively at different stages of the study (T0, T1, T2, T3, T4)	The BTT Z-score significantly correlated with the BMD Z-score at each stage, while there was variable correlation with AD-SoS Z-score at different stages of the GC treatment.
Bak-Drabik et al	2016	511	BMD_{LS} and Ad-SoS was (r=0.69) BMD_{TB} and Ad-SoS was (r=0.74)	Ad-SoS correlated significantly and positively with BMD_{LS} and BMD_{TB}
Adamczyk et al	2018	76	Ad-SoS and BMD_{TB} was (r = 0.77) Ad-SoS and BMD_{LS} was (r value range 0.42–0.69)	Ad-SoS was significantly correlated with BMD_{TB}. Ad-SoS was moderately correlated with BMD_{LS}

Table 5: Studies showing moderate correlation between QUS and DXA.

Authors	Year	Sample Size	Correlation	Comments
Brukx et al	2003	40	Correlation coefficient between QUS and DXA parameters (r = 0.14–0.50)	Moderate correlation between QUS and DXA parameters
Weeks et al	2016	389	Correlation between BUA and BMC and BMD at the femoral neck (r = 0.47, 0.49) BUA and BMC and BMD at the lumbar spine (r = 0.49, 0.50) BUA and BMC and BMD at the whole body (r = 0.54, 0.56)	BUA showed moderate correlation with BMC and BMD
Torres-Costoso et al	2018	107	Stiffness index (SI) (r = 0.43–0.52) Broadband ultrasound attenuation (BUA) (r = 0.50–0.58) Speed of sound (SOS) (r = 0.25–0.27)	Moderate correlation between BMC-DXA and QUS parameters. SI and BUA were positively correlated with DXA-measured BMC, except SOS

Table 6: Studies showing weak or no correlation between QUS and DXA.

Authors	Year	Sample Size	Correlation	Comments
Christoforidis et al	2010	26		No agreement was recorded between these 2 methods
Christoforidis et al	2011	27		No agreement was recorded between these 2 methods
Sioen et al	2011	61	Correlation between SI and BMD_{TB} was r=-0.370	SI was negatively correlated with BMD_{TB}
Srichan et al	2014	181	Correlation between BMD Z-score and SOS Z-score was (r = 0.02)	Poor correlation between BMD Z-score and SOS Z-score
Lageweg et al	2018	60	Correlation coefficient between DXA and QUS measurements was 0.291–0.462	Pearson's correlation coefficients and partial correlation coefficients presented a weak correlation between DXA and pQUS parameters
Wikiera et al	2018	43		No agreement was recorded between these 2 methods.

The positive relationship for QUS and DXA by pooled estimate, based on 13 studies, was 0.57 [95%CI: 0.46; 0.67, p-value<0.001] (**Figure 6**). The level of heterogeneity was high (I^2= 91%), and the Q-test was statistically significant (p-value<0.001).

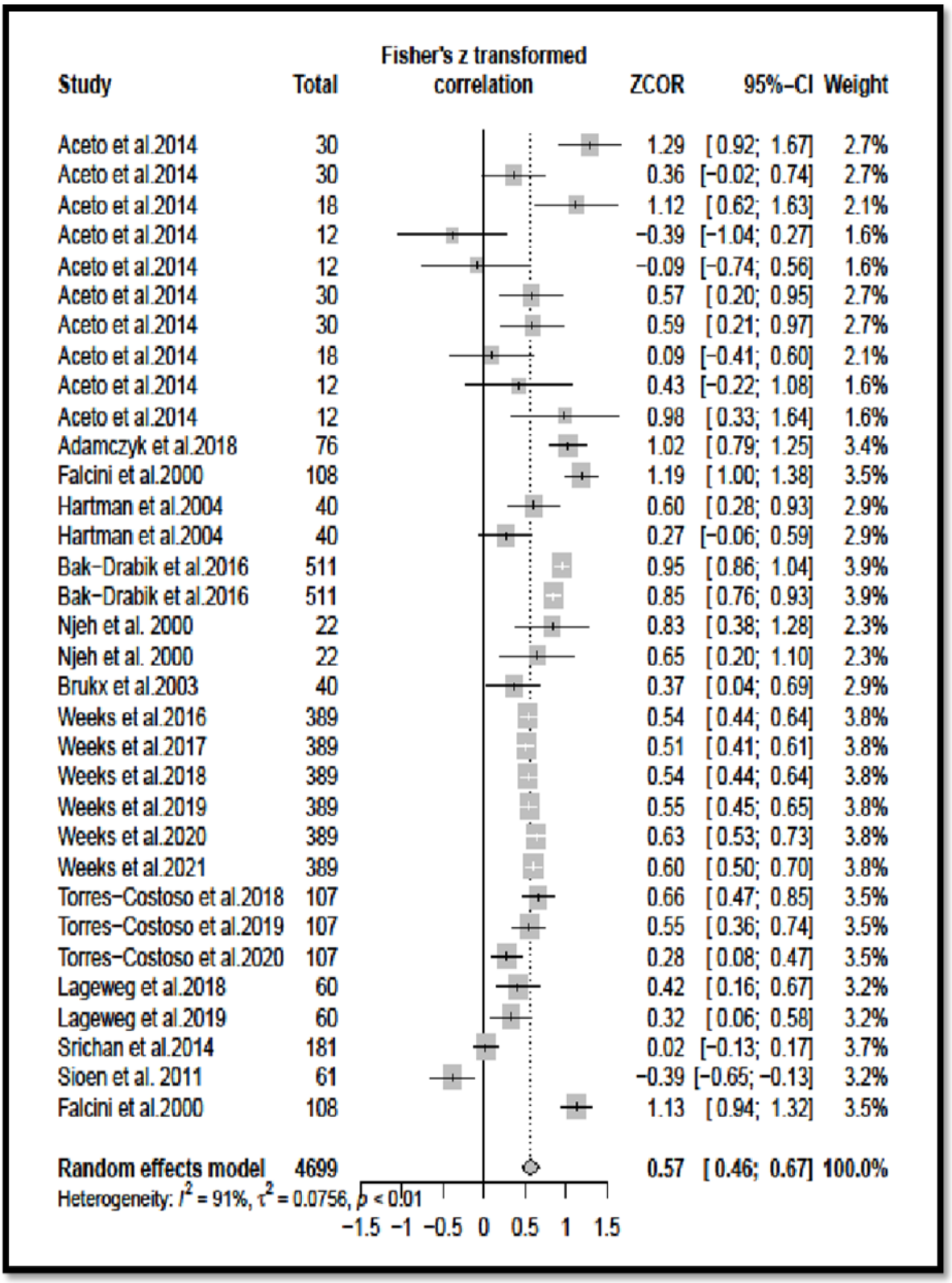

Figure 6: Forest plot of 12 studies for relationship between QUS and DXA using Fisher's transformed correlation

For the diagnosis of sensitivity, there was one potential outlier [61], which was not close to the threshold limits of eight indices. Excluding this outlier did not affect the overall estimate.

For bias publication, most studies were within the diagonal lines for the 95% confidence limits around the pooled correlation (Figure **7**).

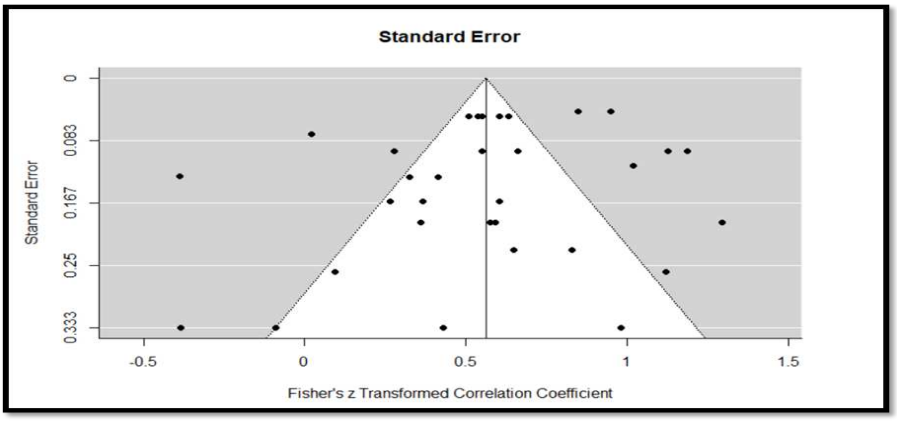

Figure 7: Funnel plot for QUS and DXA

Meta-analysis for estimating both sensitivity and specificity were based on extracting the number of TP, FP, FN and TN from three studies (**Table 7**). Two studies calculated the sensitivity and specificity of bone health index (BHI) values, comparing them with DXA parameters [28, 44] and one study compared QUS with DXA [57].

Table 7: Study-levels outcomes

Authors	Year	TP	FN	FP	TN	Sensitivity	Specificity	Weighted specificity	Weighted sensitivity
Nusman et al	2015	5.71%	8.57%	17.14%	68.57%	0.400	0.800	33.333	17.411
Neelis et al	2018	37.50%	8.33%	4.16%	50%	0.818	0.923	14.444	36.558
Lageweg et al	2018	15%	6.66%	25%	53.33%	0.692	0.681	52.222	46.031

The overall sensitivity was statistically significant (overall sensitivity= 0.71, p-value<0.001, 95% CI: 0.53 0.88] with low heterogeneity (I^2=14.8%). While the overall specificity was statistically significant (overall specificity = 0.80, p-value<0.001, 95%CI: 0.66 0.94] with slightly moderate heterogeneity (I^2= 65.5%) (**Figure 8**).

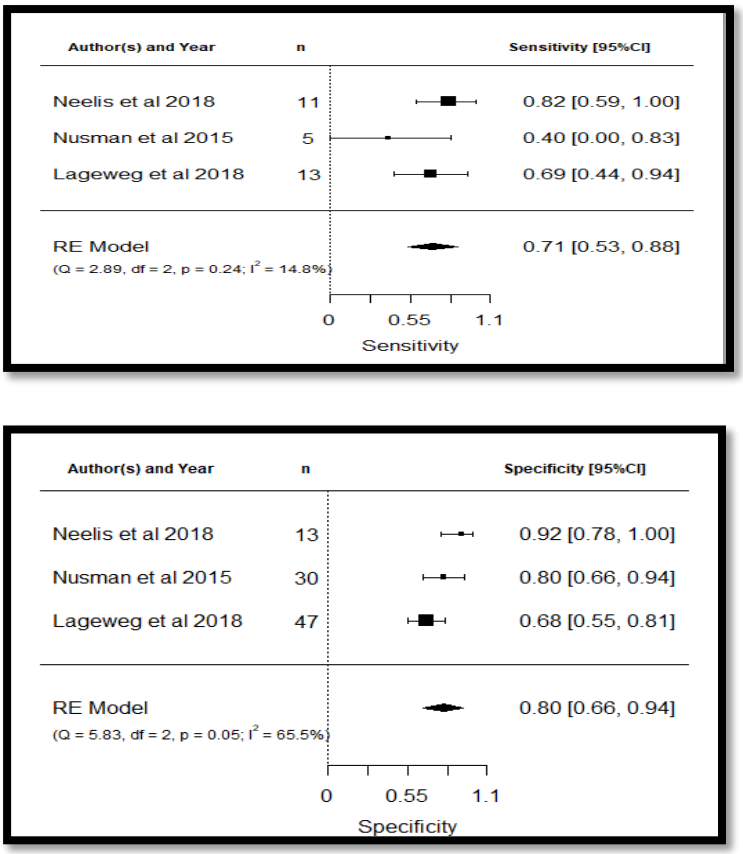

Figure 8: Forest plot for sensitivity and specificity of three studies comparing to DXA.

Discussion

This systematic review and meta-analysis determine the accuracy and correlation of different bone imaging modalities when compared to DXA, which is viewed as the most reliable and is the most widespread method for assessing bone density in children and young people.

Of the seven papers that compared digital X-ray radiogrammetry (DXR) with DXA, four found a moderate to strong degree of correlation [28, 41, 42, 45]. This suggests that DXR may provide results close to those obtained from DXA regarding measuring bone mineral density and might reflect the ability of DXR to be used as a reliable modality for evaluating bone density and for predicting fracture risk in children and young people. On the other hand, two studies concluded that the degree of correlation was poor [40] [44]. The lack of a correlation between the Z-scores derived from DXA and DXR in these studies may be due to several factors: firstly, the Z-scores of these modalities were based on different reference populations. Secondly, DXR with BHI can measure cortical thickness, metacarpal length, and width, which represent volumetric BMD (non-size dependence method), while DXA is unable to measure bone depth (it represents areal BMD). One could postulate that in fact BHI may be more reliable for assessing bone strength and predicting fracture risk in children.

Four papers compared pQCT and DXA in the measurement of bone density in healthy and unhealthy children and young people [46-49]. Only one study showed that trabecular vBMD and cortical CSA Z-scores, measured by pQCT were strongly correlated with DXA BMD Z-scores, whilst the other showed low to moderate correlation between different parameters of pQCT and DXA. The number of studies that addressed pQCT in this review, however, is

small. Although pQCT can determine and measure the total, trabecular, and cortical volumetric bone mineral density (vBMD), and provides various details about bone geometrical parameters [20], this tool has some drawbacks, such as being very sensitive to motion. It is therefore considered less suitable for use with young children. In addition, there is little pediatric reference data available in terms of different age groups, ethnicities, and range of health problems. Regarding the dose of radiation of both methods (DXA, pQCT), O'Brien et al (2018) highlighted the fact that ionizing radiation dose of DXA was significantly lower than that from pQCT, the doses being 2 μSv and 60 μSv, respectively [48]. This limits the routine use of pQCT for bone health assessment in children.

The choice of QUS as a preferred imaging modality for evaluating bone health in children and young people remains under debate. Out of a total of 17 papers, seven suggested that QUS is an acceptable technique for evaluating bone health in children and young adults with different health conditions. Conversely, the other ten papers suggested that QUS may not detect children with poor bone mineral density as compared with DXA. Although the various QUS techniques used in assessing bone status in children depend on the same principles of physics, there are many differences between these quantitative ultrasound devices in precision, accuracy, and skeletal sites that have been measured. For these reasons, use of QUS as a reliable mobility in children is still controversial.

Of the imaging modalities compared to DXA in this review, DXR appears to have the most promise, being inexpensive, widely available, and can be used easily in children because it is a quick technique with software available that allows reliable and computerized measurement of BHI. Additionally, BHI describes not only cortical thickness (T), which represents aBMD, but also measures metacarpal length and width, which represent vBMD (non-size dependence method). Furthermore, DXR is considered a safe tool for bone health in children because its radiation dose (to a peripheral site) is relatively low. The ionizing radiation dose for a PA radiograph of the hand is <0.1 μSv [DXA radiation dose (1–6 μSv)].

There are several limitations to our review. First, we only included studies published in the English language; thus, important data from studies in other languages may have been missed. Second, the analysis was limited to a population aged under 18 years. We excluded studies where the participants were both adults and children if the data relating to the children was not extractable. This decision led to the exclusion of several papers assessing HR-pQCT and MRI. Third, there existed significant statistical heterogeneity of the outcome measures due to variability across the study designs, imaging methods and parameters, follow-up duration, patient characteristics (gender, ethnicity, age, weight, height), and other qualitative attributes. Fourth, only three papers in this review demonstrated the diagnostic accuracy of the imaging modalities compared to DXA. The meta-analysis to estimate both sensitivity and specificity was based on extracting the number of TP, FP, FN, and TN from the studies. These values were only found in three studies included in this review [27,43,56]. In addition, the methodological information for conducting the imaging techniques and the time interval between the standard test (DXA) and the index test (DXR, QUS or pQCT) were not clarified in almost half of papers.

To summarize, no imaging modality can provide a full evaluation of bone health in children and young adults, with each having limitations. Our meta-analysis provides evidence for a strong correlation between DXR and DXA. Given that DXR gives a rapid, objective and 3D

evaluation of bone health, further studies are warranted to assess its use as a low cost, rapid imaging technique to assess bone parameters and predict fracture risk in children and young adults.

References

1. Bachrach, L.K. and C.M. Gordon, *Bone densitometry in children and adolescents.* Pediatrics, 2016. **138**(4): p. e20162398.
2. Consensus, A., *Consensus development conference: diagnosis, prophylaxis, and treatment of osteoporosis.* Am J Med, 1993. **94**(6): p. 646-50.
3. Jankowski, L.G., et al., *The Official Positions of the International Society for Clinical Densitometry: Cross Calibration, Least Significant Change, and Quality Assurance in Multiple Dual-Energy X-Ray Absorptiometry Scanner Environments.* Journal of clinical densitometry: the official journal of the International Society for Clinical Densitometry, 2019.
4. Shuhart, C.R., et al., *Executive summary of the 2019 ISCD position development conference on monitoring treatment, DXA cross-calibration and least significant change, spinal cord injury, peri-prosthetic and orthopedic bone health, transgender medicine, and pediatrics.* Journal of Clinical Densitometry, 2019. **22**(4): p. 453-471.
5. Bachrach, L., *Dual energy X-ray absorptiometry (DEXA) measurements of bone density and body composition: promise and pitfalls.* Journal of pediatric endocrinology & metabolism: JPEM, 2000. **13**: p. 983-988.
6. Wasserman, H., J.M. O'Donnell, and C.M. Gordon, *Use of dual energy X-ray absorptiometry in pediatric patients.* Bone, 2017. **104**: p. 84-90.
7. Bolotin, H., *DXA in vivo BMD methodology: an erroneous and misleading research and clinical gauge of bone mineral status, bone fragility, and bone remodelling.* Bone, 2007. **41**(1): p. 138-154.
8. Carter, D.R., M.L. Bouxsein, and R. Marcus, *New approaches for interpreting projected bone densitometry data.* Journal of bone and mineral research, 1992. **7**(2): p. 137-145.
9. Leonard, M.B., *Assessment of bone health in children and adolescents with cancer: promises and pitfalls of current techniques.* Medical and pediatric oncology, 2003. **41**(3): p. 198-207.
10. Crabtree, N., et al., *Diagnostic evaluation of bone densitometric size adjustment techniques in children with and without low trauma fractures.* Osteoporosis International, 2013. **24**(7): p. 2015-2024.
11. Weber, L.T. and O. Mehls, *Limitations of dual x-ray absorptiometry in children with chronic kidney disease.* Pediatric Nephrology, 2010. **25**(1): p. 3.
12. Messina, C., et al., *Pediatric dual-energy X-ray absorptiometry in clinical practice: what the clinicians need to know.* European journal of radiology, 2018. **105**: p. 153-161.
13. Khalatbari, H., L.A. Binkovitz, and M.T. Parisi, *Dual-energy X-ray absorptiometry bone densitometry in pediatrics: a practical review and update.* Pediatric Radiology, 2020: p. 1-15.
14. Weber, D.R., et al., *The utility of DXA assessment at the forearm, proximal femur, and lateral distal femur, and vertebral fracture assessment in the pediatric population: 2019 ISCD official position.* Journal of Clinical Densitometry, 2019. **22**(4): p. 567-589.
15. Hans, D., et al., *Bone microarchitecture assessed by TBS predicts osteoporotic fractures independent of bone density: the Manitoba study.* Journal of Bone and Mineral Research, 2011. **26**(11): p. 2762-2769.
16. Tong, L., J.S. Gopal-Kothandapani, and A.C. Offiah, *Feasibility of quantitative ultrasonography for the detection of metabolic bone disease in preterm infants—systematic review.* Pediatric radiology, 2018. **48**(11): p. 1537-1549.
17. McDevitt, H. and S. Ahmed, *Quantitative ultrasound assessment of bone health in the neonate.* Neonatology, 2007. **91**(1): p. 2-11.

18. Pettinato, A.A., et al., *Effects of nutrition, puberty, and gender on bone ultrasound measurements in adolescents and young adults.* Journal of adolescent health, 2006. **39**(6): p. 828-834.
19. Cheng, S., et al., *Influence of region of interest and bone size on calcaneal BMD: implications for the accuracy of quantitative ultrasound assessments at the calcaneus.* The British journal of radiology, 2002. **75**(889): p. 59-68.
20. Zemel, B.S., *Quantitative computed tomography and computed tomography in children.* Current osteoporosis reports, 2011. **9**(4): p. 284.
21. Adams, J.E., *Quantitative computed tomography.* European journal of radiology, 2009. **71**(3): p. 415-424.
22. Adams, J.E., et al., *Quantitative computer tomography in children and adolescents: the 2013 ISCD Pediatric Official Positions.* Journal of Clinical Densitometry, 2014. **17**(2): p. 258-274.
23. Damilakis, J., et al., *Radiation exposure in X-ray-based imaging techniques used in osteoporosis.* European radiology, 2010. **20**(11): p. 2707-2714.
24. Cheung, A.M., et al., *High-resolution peripheral quantitative computed tomography for the assessment of bone strength and structure: a review by the Canadian Bone Strength Working Group.* Current osteoporosis reports, 2013. **11**(2): p. 136-146.
25. Digby, M., et al., *HR-pQCT: a non-invasive 'biopsy'to assess bone structure and strength.* Archives of Disease in Childhood-Education and Practice, 2016. **101**(5): p. 268-270.
26. Pauchard, Y., et al., *Quality control for bone quality parameters affected by subject motion in high-resolution peripheral quantitative computed tomography.* Bone, 2012. **50**(6): p. 1304-1310.
27. Thodberg, H.H. and L. Sävendahl, *Validation and reference values of automated bone age determination for four ethnicities.* Academic radiology, 2010. **17**(11): p. 1425-1432.
28. Neelis, E., et al., *Bone health of children with intestinal failure measured by dual energy X-ray absorptiometry and digital X-ray radiogrammetry.* Clinical Nutrition, 2018. **37**(2): p. 687-694.
29. Page, M.J. and D. Moher, *Evaluations of the uptake and impact of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Statement and extensions: a scoping review.* Systematic reviews, 2017. **6**(1): p. 263.
30. Whiting, P.F., et al., *QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.* Annals of internal medicine, 2011. **155**(8): p. 529-536.
31. Hedges, L. and I. Olkin, *Statistical Methods for Meta-Analysis (Orlando, FL: Academic).* HedgesStatistical Methods for Meta-Analysis1985, 1985.
32. Cohen, J., *Statistical power analysis for the behavioral sciences*. 2013: Academic press.
33. Field, A.P., *The problems in using fixed-effects models of meta-analysis on real-world data.* Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences, 2003. **2**(2): p. 105-124.
34. Lewis, S. and M. Clarke, *Forest plots: trying to see the wood and the trees.* Bmj, 2001. **322**(7300): p. 1479-1480.
35. Viechtbauer, W. and M.W.L. Cheung, *Outlier and influence diagnostics for meta-analysis.* Research synthesis methods, 2010. **1**(2): p. 112-125.
36. Sani, F.M., S.A. Sarji, and M. Bilgen, *Quantitative Ultrasound Measurement of the Calcaneus in Southeast Asian Children With Thalassemia: Comparison With Dual-Energy X-ray Absorptiometry.* Journal of Ultrasound in Medicine, 2011. **30**(7): p. 883-894.
37. Njeh, C.F., et al., *Use of quantitative ultrasound to assess bone status in children with juvenile idiopathic arthritis: a pilot study.* Journal of Clinical Densitometry, 2000. **3**(3): p. 251-260.
38. Bąk-Drabik, K., et al., *Bone status assessed by quantitative ultrasound in children with inflammatory bowel disease: a comparison with DXA.* Expert review of gastroenterology & hepatology, 2016. **10**(11): p. 1305-1312.

39. Schündeln, M.M., et al., *A piece of the puzzle: the bone health index of the BoneXpert software reflects cortical bone mineral density in pediatric and adolescent patients.* PloS one, 2016. **11**(3).
40. Alshamrani, K., et al., *Estimating bone mass in children: can bone health index replace dual energy x-ray absorptiometry?* Pediatric radiology, 2019. **49**(3): p. 372-378.
41. Leijten, A.D., et al., *Can digital X-ray radiogrammetry be an alternative for dual-energy X-ray absorptiometry in the diagnosis of secondary low bone quality in children?* European journal of pediatrics, 2019. **178**(9): p. 1433-1441.
42. Mentzel, H.-J., et al., *The potential of digital X-ray radiogrammetry (DXR) in the assessment of osteopenia in children with chronic inflammatory bowel disease.* Pediatric radiology, 2006. **36**(5): p. 415-420.
43. Mulugeta, P.G., et al., *Determination of osteopenia in children on digital radiography compared with a DEXA reference standard.* Academic radiology, 2011. **18**(6): p. 722-725.
44. Nusman, C.M., et al., *Bone health of patients with juvenile idiopathic arthritis: a comparison between dual-energy X-ray absorptiometry and digital X-ray radiogrammetry.* European journal of radiology, 2015. **84**(10): p. 1999-2003.
45. van Rijn, R.R., et al., *Direct X-ray radiogrammetry versus dual-energy X-ray absorptiometry: assessment of bone density in children treated for acute lymphoblastic leukaemia and growth hormone deficiency.* Pediatric radiology, 2006. **36**(3): p. 227-232.
46. DiVasta, A., et al., *Skeletal outcomes by peripheral quantitative computed tomography and dual-energy X-ray absorptiometry in adolescent girls with anorexia nervosa.* Osteoporosis International, 2016. **27**(12): p. 3549-3558.
47. Kalkwarf, H., T. Laor, and J.A. Bean, *Fracture risk in children with a forearm injury is associated with volumetric bone density and cortical area (by peripheral QCT) and areal bone density (by DXA).* Osteoporosis international, 2011. **22**(2): p. 607-616.
48. O'Brien, C.E., et al., *Peripheral quantitative computed tomography detects differences at the radius in prepubertal children with cystic fibrosis compared to healthy controls.* PloS one, 2018. **13**(1).
49. Tsampalieros, A., et al., *Changes in trabecular bone density in incident pediatric Crohn's disease: a comparison of imaging methods.* Osteoporosis International, 2014. **25**(7): p. 1875-1883.
50. Aceto, G., et al., *Bone health in children and adolescents with steroid-sensitive nephrotic syndrome assessed by DXA and QUS.* Pediatric Nephrology, 2014. **29**(11): p. 2147-2155.
51. Adamczyk, P., M. Szczepanska, and W. Pluskiewicz, *Skeletal status assessment by quantitative ultrasound and bone densitometry in children with different renal conditions.* Osteoporosis International, 2018. **29**(12): p. 2667-2675.
52. Falcini, F., et al., *Comparison of quantitative calcaneal ultrasound and dual energy X-ray absorptiometry in the evaluation of osteoporotic risk in children with chronic rheumatic diseases.* Calcified Tissue International, 2000. **67**(1): p. 19-23.
53. Hartman, C., et al., *Assessment of osteoporosis by quantitative ultrasound versus dual energy X-ray absorptiometry in children with chronic rheumatic diseases.* The Journal of rheumatology, 2004. **31**(5): p. 981-985.
54. Brukx, L. and J. Waelkens, *Evaluation of the usefulness of a quantitative ultrasound device in screening of bone mineral density in children.* Annals of human biology, 2003. **30**(3): p. 304-315.
55. Weeks, B., et al., *Is calcaneal broadband ultrasound attenuation a valid index of dual-energy x-ray absorptiometry-derived bone mass in children?* Bone & joint research, 2016. **5**(11): p. 538-543.
56. Torres-Costoso, A., et al., *Agreement Between Dual-Energy X-Ray Absorptiometry and Quantitative Ultrasound to Evaluate Bone Health in Adolescents: The PRO-BONE Study.* Pediatric exercise science, 2018. **30**(4): p. 466-473.

57. Lageweg, C.M., et al., *Evaluation of bone mineralization in former preterm born children: Phalangeal quantitative ultrasound cannot replace dual-energy X-ray absorptiometry.* Bone reports, 2018. **8**: p. 38-45.
58. Srichan, W., et al., *Bone status measured by quantitative ultrasound: a comparison with DXA in Thai children.* European journal of clinical nutrition, 2016. **70**(8): p. 894-897.
59. Wikiera, B., et al., *The assessment of skeletal status in young patients with Turner syndrome by 2 densitometric techniques: Phalangeal quantitative ultrasound and dual energy X-ray absorptiometry.* Advances in clinical and experimental medicine: official organ Wroclaw Medical University, 2018. **27**(6): p. 759-764.
60. Christoforidis, A., et al., *Comparative study of dual energy X-ray absorptiometry and quantitative ultrasonography with the use of biochemical markers of bone turnover in boys with haemophilia.* Haemophilia, 2011. **17**(1): p. e217-e222.
61. Sioen, I., et al., *The relationship between paediatric calcaneal quantitative ultrasound measurements and dual energy X-ray absorptiometry (DXA) and DXA with laser (DXL) as well as body composition.* International journal of obesity, 2011. **35**(1): p. S125-S130.
62. Christoforidis, A., et al., *Bone status of children with hemophilia A assessed with quantitative ultrasound sonography (QUS) and dual energy X-ray absorptiometry (DXA).* Journal of pediatric hematology/oncology, 2010. **32**(7): p. e259-e263.