Bone 92 (2016) 58–69

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Review Article

Diagnostic devices for osteoporosis in the general population: A systematic review

M.P. Høiberg [a,b,*], K.H. Rubin [c], A.P. Hermann [d], K. Brixen [d], B. Abrahamsen [c,e]

[a] Department of Research, Hospital of Southern Norway, Kristiansand, Norway
[b] Department of Clinical Research, University of Southern Denmark, Odense, Denmark
[c] OPEN, Odense Patient Data Explorative Network, Department of Clinical Research, University of Southern Denmark, Odense University hospital, Denmark
[d] Department of Medical Endocrinology, Odense University Hospital, Odense, Denmark
[e] Department of Medicine, Holbæk Hospital, Holbæk, Denmark

ARTICLE INFO

Article history:
Received 6 July 2016
Revised 9 August 2016
Accepted 14 August 2016
Available online 16 August 2016

Keywords:
DXA
Quantitative ultrasound
Quantitative computed tomography
Trabecular bone score
Population-based
QUADAS 2

ABSTRACT

Introduction: A diagnostic gap exists in the current dual photon X-ray absorptiometry (DXA) based diagnostic approach to osteoporosis. Other diagnostic devices have been developed, but no comprehensive review concerning the applicability of these diagnostic devices for population-based screening have been performed.
Material and methods: A systematic review of Embase, Medline and the Cochrane Central Register for Controlled Trials was performed for population-based studies that focused on technical methods that could either indicate bone mineral density (BMD) by DXA, substitute for DXA in prediction of fracture risk, or that could have an incremental value in fracture prediction in addition to DXA. Quality of included studies was rated by QUADAS 2.
Results: Many other technical devices have been tested in a population-based setting. Five studies aiming to indicate BMD and 17 studies aiming to predict fractures were found. Overall, the latter studies had higher methodological quality. The highest number of studies was found for quantitative ultrasound (QUS). The ability to indicate BMD or predict fractures was moderate to minor for all examined devices, using reported area under the curve (AUC) of Receiver Operating Characteristic curves values as standard.
Conclusions: Of the methods assessed, only QUS appears capable of perhaps replacing DXA as standalone examination in the future whilst radiographic absorptiometry could provide important information in areas with scarcity of DXA. QUS may be of added value even after DXA has been performed. Evaluation of proposed cutoff-values from population-based studies in separate population-based cohorts is still lacking for most examination devices.

Contents

* Corresponding author at: Department of Research, Hospital of Southern Norway, Kristiansand, Egsveien 4, N-4604 Kristiansand, Norway.
E-mail addresses: mikkelhoiberg@gmail.com (M.P. Høiberg), krubin@health.sdu.dk (K.H. Rubin), pernille.hermann@rsyd.dk (A.P. Hermann), kbrixen@health.sdu.dk (K. Brixen), b.abrahamsen@physician.dk (B. Abrahamsen).

http://dx.doi.org/10.1016/j.bone.2016.08.011
8756-3282/

1. Introduction

Osteoporosis is characterized by low bone mass and micro-architectural deterioration of bone tissue leading to increased risk of fractures [1]. One in three women and one in six men will suffer at least one osteoporotic fracture during their lifetime [2]. The disease is silent until the event of a fracture. Cross-sectional studies have consistently found that osteoporosis is under-diagnosed both in the general population [3] and in high-risk groups [4,5]. The diagnostic gap is further exaggerated by the fact that even in industrialized countries such as the US, examination rates for osteoporosis and prescription of pharmacological therapy for osteoporosis have been declining [6]. Even though Fracure Liasion Services are cost-effective and most likely also cost-saving [7], these coordinated services are lacking in many institutions.

The current gold standard for diagnosing osteoporosis is dual photon X-ray absorptiometry (DXA), defined as bone mineral density (BMD) >2.5 standard deviations below the mean of young reference populations [8]. Fracture risk is inversely related to BMD [9], but as BMD is normally distributed, a much larger proportion of the population has osteopenic BMD values (i.e. in the range from −1 to −2.5 SD) rather than values in the osteoporotic range. Therefore, the majority of fractures occur in the osteopenic group despite a lower individual risk of fracture [10]. Furthermore, the elements of altered bone quality and individual risk of falling, both of which have substantial impact on the risk of fractures [11,12], are not captured by DXA, thus potentially reducing the predictive value of an examination. Finally, DXA is not universally available in all countries or healthcare systems. Therefore, the search for other tools capable of either differentiating patients at high risk of future osteoporotic fractures or alternatively identifying the same high risk groups as DXA by employing simpler, more accessible methods is ongoing and could help narrow the diagnostic gap in osteoporosis. Previously, we performed a systematic review of the fracture risk prediction of different algorithms based on clinical risk factors, finding that more complex algorithms did not perform better than simpler ones [13]. Recently, reviews of selected technical equipment designed to be used as an alternative or adjunct to DXA have been performed [14–16]. Though interesting, most comparisons of technical devices against DXA have been performed in non-population based settings, or by combining population-based with non population-based studies. This poses a problem as diagnostic performance is critically dependent on the disease prevalence in the evaluated population. Furthermore, different populations have been used as reference populations amongst manufactures of the same equipment, leading to differences in reported standard deviations [17]. This has further added to the confusion and made the search for a common diagnostic T-score cutoff an act of futility. Finally, the correlation between methodologies cannot be assumed to be independent of artifacts such as osteoarthrosis that may have an uneven influence between devices and between measurement sites. Thus, the utility of technical devices for population-based screening for osteoporosis has not been thoroughly addressed.

We aim to provide a systematic review of existing technical methods to indicate DXA-defined osteoporosis, tested in a population-based setting using a properly documented methodology. Specifically we wish to answer three questions: Which technical methods may indicate DXA defined osteoporosis? Which technical methods may substitute for DXA for the prediction of fracture risk? Furthermore, which technical

methods for osteoporosis have incremental value in fracture risk assessment if BMD by DXA is already known? The review evaluates both the performance of the devices and the quality of the evidence base supporting their clinical use in osteoporosis diagnosis and risk assessment.

2. Materials and methods

The Population, Intervention, Comparison and Outcome and Study Design (PICOS) method was used to define research questions and search strategy [18].

2.1. Studies

We included randomized controlled trials (RCTs), including cluster RCTs, controlled (non-randomized) clinical trials (CCTs) and cluster trials as well as prospective, cross-sectional and retrospective cohort studies concerning humans above the age of 40 years from a population-based setting. Included studies should state the proportion of subjects within the defined study population with known osteoporosis and the basis of this diagnosis. Thus, an adequate description of the examined population in included studies was mandatory.

2.2. Interventions

All diagnostic approaches relying on devices other than biochemistry or the sole use of clinical risk scores were evaluated.

2.3. Comparison

Comparison with the current gold standard DXA or focusing on the prediction of fractures as endpoint was mandatory.

2.4. Outcomes

Relevant outcomes were either prospective identification or stratification of persons at high risk of incident fractures, alternatively identification of persons shown by DXA to have osteoporosis. Included studies had to provide information necessary for the calculation of performance characteristics in the form of sensitivity and specificity or ROC curves over a spectrum of different values of sensitivity and specificity for the prediction of fractures or the prediction of DXA-diagnosed osteoporosis. Authors of papers where such data were missing were *not* contacted for supplementary information. Studies based on hypothetical cohorts and focusing only on cost-effectiveness analyses were excluded.

Potentially eligible papers were assessed by two reviewers (MPH and KHR). Any papers with discrepant evaluations for eligibility in the screening process by the two reviewers were discussed in the group to establish consensus.

2.5. Search strategy and flow

Literature search was conducted in Embase (OVID interface from 1980 and onwards), Medline (OVID interface from 1948 and onwards) and The Cochrane Central Register for Controlled Trials to identify relevant publications up to the search date of October 28, 2014, supplemented with an additional search in Pubmed to include papers published within the last four months that had yet to be indexed properly in Embase and Medline. The databases were searched for papers including the search-terms osteoporosis AND (screening OR comparison OR prediction/predictive) AND one or more examination modalities in question, using truncated search terms (see Appendix 1 for exhaustive search strategy). Paper types were restricted to "article, randomized controlled trial, case report, clinical trial, comparative study, controlled clinical trial, observational study, pragmatic clinical trial, twin study or validation studies" in Medline. Papers in other languages than English, Nordic languages or German were excluded due to the inability of the authors to interpret papers in full text. Reference lists of included studies and other relevant reviews identified through the search were also screened for relevant literature missed during the initial search.

In order to rule out bias due to duplicate publications, included studies were evaluated for the existence of publications from the same cohort. In case of multiple duplications analyzing the same variable in the same population, the publication using the largest proportion of available study subjects from the cohort in question was used.

2.6. Assessment of methodological quality and data abstraction

Information about study design, inclusion- and exclusion criteria, number of participants, setting, baseline examinations, follow-up and data-collection as well as results were recorded. Methodological quality of included studies was assessed using the QUality Assessment tool for Diagnostic Accuracy Studies (QUADAS) 2 checklist [19], as recommended by the Cochrane Handbook of Systematic Reviews [20]. In papers reporting both developmental and validation cohorts for quality assessment, the cohorts were pooled when assessing the methodological quality. The checklists were employed independently by MPH and KHR and discrepancies in rating scores were discussed for consensus. Transcripts of QUADAS checklists for studies with comparison with DXA or for prediction of fractures are found in appendix 2. As seen, QUADAS 2 were slightly modified due to the nature of the research questions. Thus, for studies aiming at indication of BMD results, 1 item (signaling question 1) was slightly modified whilst item 2, 10 and 17 were found irrelevant (shown in gray tones in Appendix 2). Three items (item 4, 18 and 20) were added to the QUADAS 2 checklist. For studies aiming at prediction of fractures, item 1 was also rephrased, and item 2, 10, 16 and 17 were found irrelevant and dropped (shown in gray tones). Again, items 4, 18 and 20 were added to the checklist.

First, studies were deemed population based on the QUADAS signaling question 6 ("Is there concern that the included patients do not match the review question"), which summarizes that participants should be recruited from the general population in an unselected matter, with description of inclusion- and exclusion criteria and avoiding inappropriate exclusions. Studies that included random samples of participants from nationwide central person registries, household-, telephone-, voter- or resident-listings were categorized as "population-based with certainty". Similarly, random samples from health insurance databases or from multiple geographic locations of the same country were regarded as "population-based with certainty". In contrast, studies including participants from secondary care centers, recruited at health fairs, by local advertisement, or included on the basis of previous examinations or the existence of known risk factors were deemed "not population-based with certainty" and therefore not included in the analysis.

3. Results

3.1. Identification of studies

A flowchart of the papers evaluated is shown in Fig. 1. A total of 3776 non-duplicated papers were found in Embase and Medline and supplemented by 635 in Pubmed using the abovementioned search strategy. The papers were screened in a two-step manner. Firstly, all titles were examined for possible relevance. If irrelevance could not be ruled out with certainty on title alone, papers were screened using the abstract. In total, 4242 were found irrelevant with certainty on the basis of title or abstract. A hand-search found additional 14 papers, leaving 183 publications for full text review. MPH identified 39 studies of interest fulfilling the PICO criteria listed above and included in the initial review.

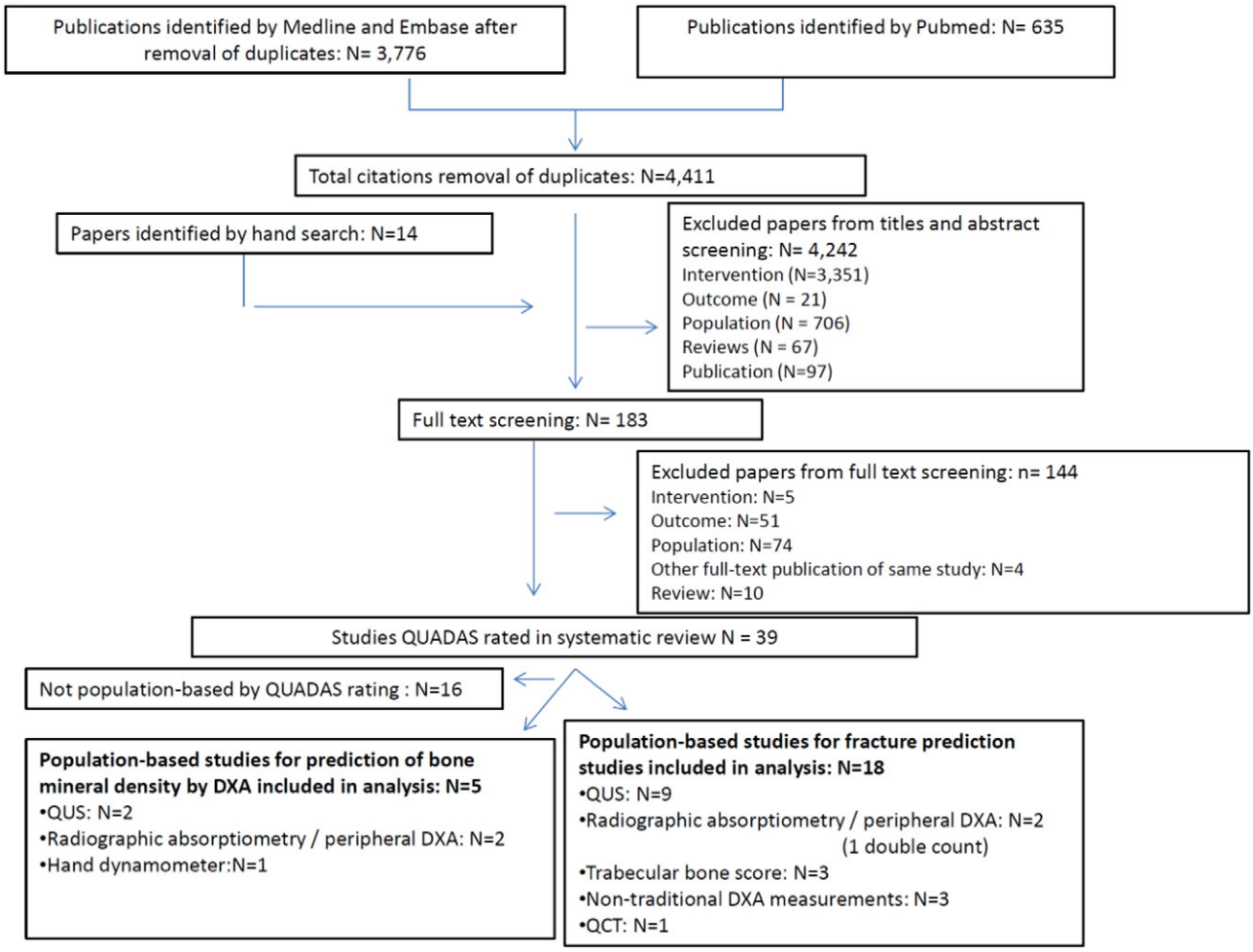

Fig. 1. Flowchart of included papers.

3.2. General characteristics of included studies to predict osteoporosis

A total of 15 different studies tested the ability of other technical devices to indicate osteoporosis as defined by DXA, supplying sensitivity and specificity or an overall ROC curve AUC value. All of these were cross-sectional studies. Nine of these studies used various devices based on the technique known as quantitative ultrasound (QUS), three used radiographic absorptiometry, two employed dental panoramic radiogram and one used a hand dynamometer. The QUADAS 2 ratings of the individual studies is shown in Table 1, whilst Fig. 2 illustrates the distribution of overall methodological quality of rated studies as evaluated with the QUADAS 2 checklist. Rated studies complied with an average of 8.5 QUADAS 2 items (range 5–11). Five of the rated studies fulfilled QUADAS 2 criteria for being "population-based with certainty" (score 1 on QUADAS 2 cumulative item 6), whilst the other ten studies were found to be not population-based (gray color tone in Table 1) and thus dropped from further analysis as shown in Fig. 1. Table 2 presents an overview of study details on these studies, including a total of 1731 women and 363 men. None of the studies included >1000 participants. A total of 7–42% of participants had osteoporotic BMD values by DXA, primarily depending upon method of inclusion and age-range of participants.

3.3. General characteristics of included studies to predict fractures

A total of 24 different studies tested the ability of technical devices to predict fractures as an outcome measure, supplying sensitivity and specificity or an overall ROC curve AUC value. 23 studies used prospective fracture identification and one used a retrospective/cross-sectional perspective [21]. Six of these studies employed either multiple DXA measurements or combinations of DXA results with other measurements or non-conventional cutoff-criteria for defining high risk populations. Five studies included trabecular bone score in the fracture prediction, nine studies employed QUS, four studies used radiographic absorptiometry (one study employing several measures) and one study used Quantitative computed tomography (QCT) for fracture prediction. Fig. 3 shows the distribution of overall methodological quality of rated studies as evaluated with the QUADAS 2 checklist, whilst Table 3 shows the QUADAS 2 ratings of the individual studies. On average, the studies complied with 12.5 QUADAS 2 items (range 7–16). The methodological quality of studies aiming at prediction of fractures were thus on average better than studies aiming at prediction of BMD by DXA. Eighteen of the rated 24 studies fulfilled the QUADAS criteria for being "population-based with certainty" (score 1 on QUADAS cumulative item 6). An overview of these studies is presented in Table 4. A total of 198,739 women and 17,033 men aged 45 or above (apart from one study including participants from the age of 18 or above) participated in these studies. Twelve of these 18 studies included >1000 participants.

3.4. Overall quality and effect measure of included studies

In total, 23 studies fulfilling QUADAS criteria for being population based were included for further review.

Table 1
QUADAS 2 rating of studies using other diagnostic methods to indicate DXA-defined osteoporosis (see Appendix 2 or Fig. 2 for Item-definitions in QUADAS). Studies with overall score of 1 in cumulative item 6 were deemed population-based with certainty. Non-included studies (score 2 or 3 on item 6) are shown in gray color tone [60–69].

	QUADAS item no:																
First Author (reference)	1	3	4	5	6	7	8	9	11	12	13	14	15	16	18	19	20
Hand Dynamometer																	
Kärkkäinen [37]	1	1	1	1	1	3	2	2	1	3	1	3	3	1	2	2	1
Dental Panoramic Radiograph																	
Leite [60]	3	1	3	1	3	1	1	1	1	1	1	1	1	3	2	2	2
Karayianni [61]	2	1	1	3	2	3	1	3	1	3	1	3	3	1	1	3	2
Quantitative ultrasound																	
Massie [62]	1	1	3	3	3	3	3	3	1	3	1	3	3	1	1	3	2
Naganathan [63]	3	1	3	3	3	3	1	3	1	3	1	3	1	1	1	1	2
Nairus [64]	3	1	3	3	3	3	2	2	1	3	1	3	3	1	1	3	2
Ayers [65]	3	1	3	3	2	3	1	3	1	3	1	3	3	1	1	3	2
Kim [66]	3	1	3	3	3	1	1	1	1	1	1	1	3	1	1	1	2
Ikeda [22]	1	1	1	1	1	3	2	2	1	3	1	3	1	1	1	1	2
Kung [67]	3	1	1	1	3	3	1	3	1	3	1	3	3	1	1	3	2
Gudmundsdottir [23]	1	1	1	1	1	2	1	2	1	2	1	2	3	1	1	3	1
Kung [68]	3	1	1	1	3	3	1	3	1	3	1	3	3	1	1	3	1
Radiographic absorptiometry																	
Hansen [33]	1	1	1	1	1	3	2	2	1	3	1	3	3	1	1	3	1
Gasser [69]	2	1	1	2	2	3	2	2	1	3	1	3	3	1	1	3	2
Lekamwasam [32]	1	1	1	1	1	1	2	2	1	3	1	3	3	1	1	3	2

Fig. 4 shows the reported performance of the examined diagnostic devices as well as the quality (QUADAS rating) of the included studies. The size of the bubbles reflects the size of the study. On an aggregate level, there seems to be a trend towards higher quality studies reporting lower levels of performance overall. Apart from the QCT, no method seems to perform significantly better than other methods.

General characteristics of 23 population based studies using technical devices other than DXA to predict osteoporosis or fractures are discussed below according to device type.

3.5. Quantitative ultrasound (QUS)

3.5.1. Diagnostic use: QUS for indication of BMD by DXA

A total of two studies using QUS measures to indicate DXA results in a population-based setting were identified [22,23], including a total of total of 831 women and 145 men (Table 2). Table 1 summarizes our QUADAS rating for the quality of these studies. As shown in Table 2, included participants had a DXA-defined osteoporosis prevalence of 27–38% amongst women and 15% amongst men. These smaller population-based studies of QUS for prediction of BMD found AUC values of 0.71–0.76 for the indication of osteoporosis by DXA. None of the studies tested their cutoff-values in independent population-based samples.

3.5.2. Prognostic use: QUS for prediction of fractures

One cross-sectional [21] and eight prospective ([10,24–30] population-based studies focusing on incident fractures were found including 34,172 women and 5607 men (Tables 3 and 4).

Two studies [24,26] described optimized sensitivities and specificities of applied QUS examinations, both studies optimizing fracture prediction by a combination of QUS and knowledge about risk factors (Table 4). With 5.1% incident fractures over 2.8 yrs, the study by Guessous et al. [24] included participants with roughly double the risk of fractures as compared to 3.9% fracture incidence over 4.0 yrs in the study by Dargent-Molina et al. [26], even though the included age-groups were similar. Using Youdens index [31] as a measure of the discriminatory models chosen by Guessous et al. [24] and Dargent-Molina et al. [26], values of 0.16 and 0.33 could be calculated from reported figures. Amongst the seven studies using AUC values, the study by Durosier et al. [27] reported higher AUC values than the other studies. However, here QUS measurements were combined with either age alone or several clinical risk factors, potentially increasing the discriminatory value of the proposed model. The six other studies reported AUC values of 0.62–0.73 despite differences in design (one cross-sectional study, five prospective studies) and risk of fractures (2.9% yearly in the OST-PRE study [29] compared to 1.0% in the studies by Bauer et al. [25] and Stewart et al. [28]). In the study by Durosier et al. [27], the discriminatory value of the QUS-derived stiffness index combined with age was not superior to the model including information about risk factors only.

3.5.3. Quantitative ultrasound for fracture risk assessment when BMD by central DXA is already known

No prospective studies had examined the effect of QUS examination after results of DXA were known. However, regression analyses in most

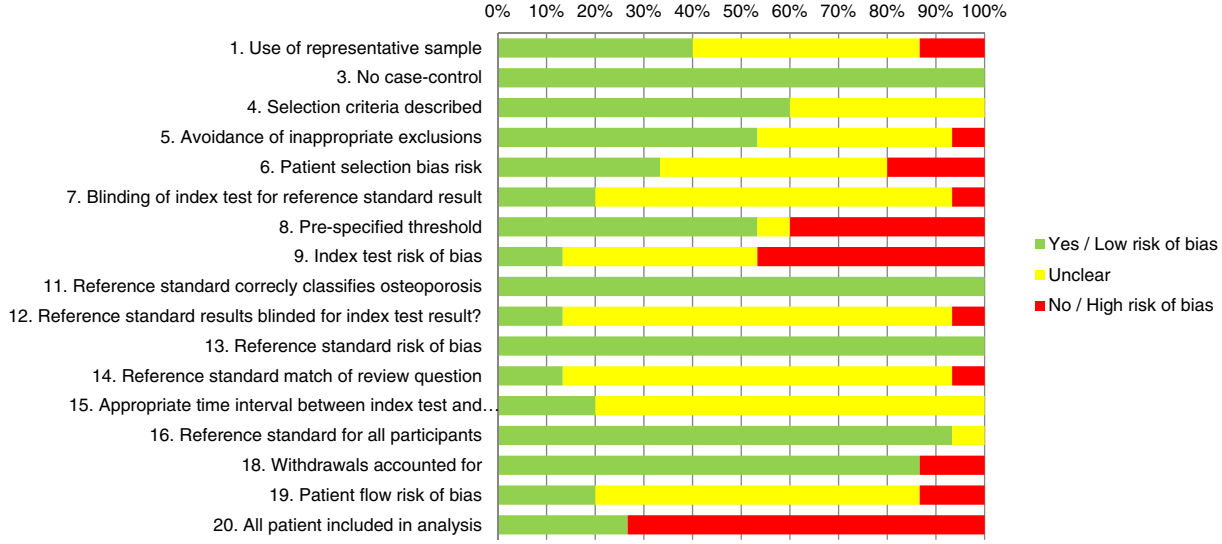

Fig. 2. Methodological quality of studies ($N = 15$) for the indication of bone mineral density by osteodensitometry as assessed by QUADAS 2 checklist.

Table 2
Population based studies using other diagnostic methods to indicate DXA-defined osteoporosis.
HD: Hand dynamometer QUS: Quantitative ultrasound RA: Radiographic absorptiometry NR: not reported CS: Cross-sectional.

Technical device	First author	Study design	Recruitment method	Enrolled women/men; total n	Proportion with osteoporosis at baseline (%)	Age (years), range	AUC, mean (range)
HD	Kärkkäinen [37]	CS	Subsample of volunteers to postal survey	750/0	11	65+	0.76 (NR)
QUS	Gudmundsdottir [23]	CS	Age-stratified random sample from population-register	172/0	27	70–85	0.76 (NR)
QUS	Gudmundsdottir [23]	CS	Age-stratified random sample from population-register	0/145	15	70–85	0.75 (NR)
QUS	Ikeda [22]	CS	Randomly selected from resident register in two regions	659/0	38.5	20–79	0.712 (NR)
RA	Hansen [33]	CS	Consecutive subgroup from register-based random sample	0/218	6.9	60–74	0.75 (0.58–0.86)
RA	Lekamwasam [32]	CS	Volunteers to public announcements	150/0	42	48+	0.76 (0.62–0.90)

[26,28–30] but not all [25] included studies find that information derived from QUS and DXA are additive and independent of the order of investigation.

3.6. Phalangeal radiographic absorptiometry (RA) and forearm peripheral DXA

3.6.1. Diagnostic use: phalangeal radiographic absorptiometry or forearm peripheral DXA for indication of BMD by DXA

Two studies evaluating the ability of phalangeal radiographic absorptiometry to indicate osteoporosis in a population-based setting were found. They included 150 women [32] and 218 men [33], respectively. A total of 42% of included women and 6.9% of included men had osteoporosis by central DXA. Models were optimized for discriminatory ability and not tested in separate samples. Reported AUC values in the two studies were similar around 0.75 – but both studies reported wide confidence intervals due to their small size.

3.7. Prognostic use: phalangeal radiographic absorptiometry or forearm peripheral DXA for prediction of fractures

Two population-based prospective studies including a total of 70,329 women and 5206 men were found [10,34]. Mean reported AUC values for fracture discrimination were 0.64–0.71. In the study by Friis-Holmberg et al. [34], fracture discrimination increased significantly by the addition of data on risk factors as summarized in the FRAX risk score.

3.7.1. Phalangeal radiographic absorptiometry or peripheral DXA for fracture risk assessment when BMD by central DXA is already known

No studies had examined the ability of the device to predict fractures as an add-on examination after central DXA.

3.8. Calcaneal single or dual X-ray absorptiometry

3.8.1. Diagnostic use: calcaneal single or dual X-ray absorptiometry for indication of BMD by DXA

No population-based studies testing this hypothesis were found.

3.8.2. Prognostic use: calcaneal single or dual X-ray absorptiometry for the prediction of fractures

A single population-based study was found, evaluating a total of 79,185 women with the Norland Osteoanalyzer Single X-ray absorptiometry of the heel [10]. A total of 11.0% reported previous fractures, 6.4% had osteoporotic DXA values. ROC curves for prediction of fractures during one year of follow-up gave an AUC value of 0.67 (SE not reported).

3.8.3. Phalangeal radiographic absorptiometry or peripheral DXA for fracture risk assessment when BMD by central DXA is already known

No population-based studies testing this hypothesis were found.

3.9. Quantitative computed tomography (QCT)

3.9.1. Diagnostic use: QCT for indication of BMD by DXA

No studies examining this research question were identified.

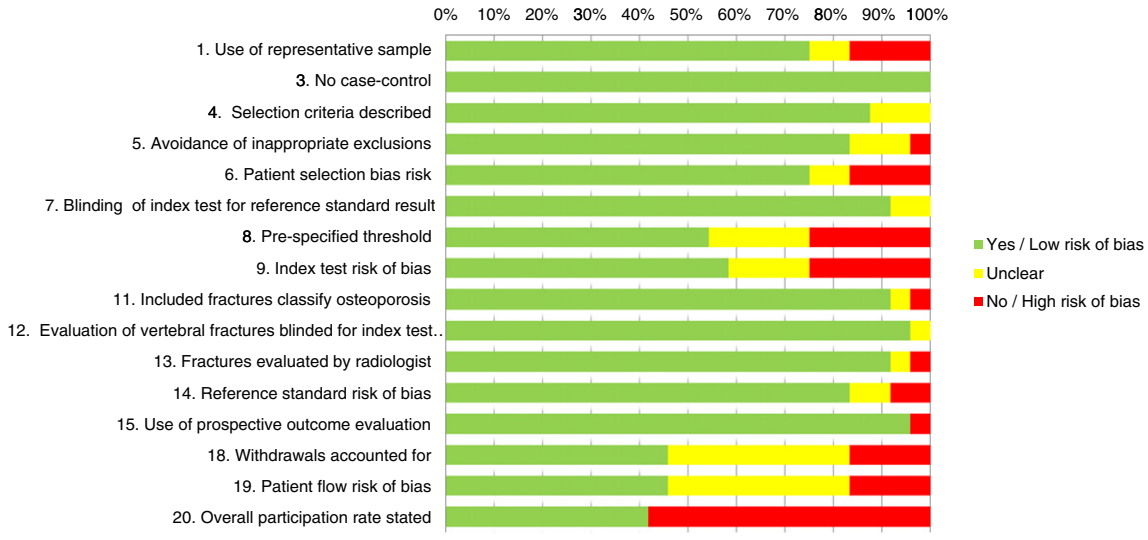

Fig. 3. Methodological quality of studies ($N = 24$) for the prediction of fractures as assessed by QUADAS 2 checklist.

Table 3
QUADAS rating of based studies using other diagnostic methods than DXA to predict fractures (see Appendix 2 or Fig. 3 for Item-definitions in QUADAS). Studies with overall score of 1 in cumulative item 6 were deemed population-based with certainty. Non-included studies (score 2 or 3 on item 6) are shown in gray color tone [70–75].

	QUADAS II item no:															
First Author (reference)	1	3	4	5	6	7	8	9	11	12	13	14	15	18	19	20
Peripheral DXA																
Barr [70]	3	1	1	1	3	1	1	1	1	1	2	2	1	2	2	2
Brismar [71]	2	1	1	2	2	1	1	1	1	1	1	1	1	1	1	2
Alternative DXA evaluation																
Leslie [72]	2	1	3	3	2	1	1	1	1	1	1	1	1	3	3	2
Wu [43]	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	2
Leslie [46]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Berry [44]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Frost [45]	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	1
Quantitative Computed Tomography																
Black [35]	1	1	1	1	1	1	3	3	1	1	1	1	1	3	3	1
Quantitative ultrasound																
Guessous [24]	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1
Bauer [25]	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	1
Dargent–Molina [26]	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	2
Durosier [27]	1	1	1	1	1	1	1	1	1	1	1	1	1	3	3	2
Miller [10]	1	1	1	1	1	1	1	1	1	1	3	3	1	3	3	1
Glüer [21]	1	1	1	1	1	3	2	2	2	1	1	2	2	2	2	2
Kwok [73]	3	1	1	1	3	1	2	2	1	1	1	1	1	1	1	2
Stewart [28]	1	1	1	1	1	1	3	3	1	1	1	1	1	1	1	1
Huopio [29]	1	1	1	1	1	1	3	3	3	1	1	3	1	1	1	1
Hollaender [30]	1	1	1	1	1	1	3	3	1	1	1	1	1	1	1	2
Radiographic absorptiometry																
Friis–Holmberg [34]	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Trabecular bone score																
Briot [40]	1	1	1	1	1	1	2	2	1	1	1	1	1	2	2	2
Leslie [74]	2	1	3	3	2	1	2	2	1	1	1	1	1	1	1	2
Boutroy (41)	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	2
Iki (42)	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	2
Hans [75]	2	1	3	3	2	1	1	1	1	1	1	1	1	3	3	2

3.9.2. Prognostic use: QCT for prediction of fractures and QCT for fracture risk assessment when BMD by central DXA is already known

Only a single study evaluating the use of QCT in a population-based setting was uncovered during the literature review, including 3347 men from the MrOS study [35]. No single QCT parameter had higher discriminatory value (HR per SD decrease) than areal BMD by DXA, but the QCT variables percent cortical volume and cross-sectional area of the femoral neck remained predictors of future fracture after adjustment for areal BMD. The ability to discriminate individuals with fractures from individuals without fractures were similar between the two methods of examination (AUC 0.85 for DXA alone and 0.86 for QCT alone (SE not reported)). Overall, QCT did not have any additive effect on the prediction of fractures (AUC values of 0.86 for QCT plus DXA (ns)).

3.10. HR-pQCT

3.10.1. Diagnostic use: HR-pQCT for indication of BMD by DXA

A single smaller ($n = 72$) population-based study examining the correlation between DXA derived cortical index and HR-pQCT-derived volumetric index was found, showing good correlation between these indices ($R = 0.798$) [36]. However, this study did not focus on the prediction of BMD defined osteoporosis and was therefore not included in the review. Thus, no population-based studies examining this research question was uncovered during the literature review.

3.10.2. Prognostic use: HR-pQCT for prediction of fractures

No population-based studies examining this research question was uncovered during the literature review.

3.10.3. HR-pQCT for fracture risk assessment when BMD by DXA is already known

No population-based studies examining this research question was uncovered during the literature review.

3.11. Dental Panoramic Radiographs (PR)

3.11.1. Diagnostic use: PR for indication of BMD by DXA

No population-based studies that focused on prediction of BMD were found in the literature review.

3.11.2. Prognostic use: PR for the prediction of fracture risk

No population-based studies were found that focused on prediction of fracture risk.

3.11.3. PR for fracture risk assessment when BMD by central DXA is already known

No studies focusing on this issue were found during the literature review.

3.12. Hand dynamometer

3.12.1. Diagnostic use: hand dynamometer for indication of BMD by DXA

A single population-based study including 750 women was identified [37], evaluating the use of a hand dynamometer for prediction of osteoporosis by BMD. Using a post-hoc defined cutoff for grip strength of the dominant hand, prediction of osteoporosis with an AUC of 0.76 was found (SE not reported).

3.12.2. Prognostic use: hand dynamometer for the prediction of fracture risk

In the above-mentioned study, low grip strength was also a significant predictor for future fractures during 2.9 yrs of follow-up; however the publication did not reveal any values for the fracture prediction [37].

3.12.3. Hand dynamometer for fracture risk assessment when BMD by central DXA is already known

No studies examining this research question was uncovered during the literature review.

3.13. Trabecular bone score (TBS)

TBS evaluates pixel gray-level variations of a lumbar spine DXA image [16]. It addresses the problem that fracture risk of certain subgroups as patients with diabetes [38] or primary hyperparathyroidism [39] does not seem to be adequately captured by BMD alone.

3.13.1. Diagnostic use: TBS for indication of BMD by DXA

The notion of using TBS for prediction of BMD does not make sense as TBS is based on information collected during a normal DXA. No population-based studies that focused on indicating BMD were found.

3.13.2. Prognostic use: TBS for prediction of fracture risk and TBS for fracture risk assessment when BMD by central DXA is already known

Three studies were found to be population-based with certainty, including 3634 women [40–42]. The predictive ability of TBS was 0.60–0.68 in these studies, all studies with comparable discriminatory ability

Table 4
Population based studies using other diagnostic methods than DXA to predict fractures.
DXA: dual X-ray Absorptiometry. pDXA: peripheral DXA.QCT: Quantitative computed tomography. QUS: quantitative ultrasound. RA: radiographic absorptiometry TBS.: Trabecular bone score. NR: not reported. CS: cross-sectional P: prospective NA: not applicable.

Technical device	Main author	Study design	Recruitment method	Enrolled women/men; total n	Follow-up/yrs	Proportion with osteoporosis/event of interest (%)	Age (years), range	AUC, mean (range)
DXA	Wu [43]	P	Random samples of women from voter registration lists in Oregon, Baltimore and Pennsylvania, at their second visit	4948/0	12	NR/30.2	67–74	NR
DXA T-score offset	Leslie [46]	P	Random selection from residential phone numbers	4575/1813	10.1	NR/10.3	50+	0.69 (0.67–0.71)
Heel SXA	Miller [10]	P	Randomly chosen names of women at 4236 GP-offices	79,185/0	1	6.4/1.6	50+	0.67 (NR)
Multiple DXA	Frost [45]	P	Community mailings to all aged >60 yrs without osteoporosis in one Australian community	1008/0	7.1	0/20.5	60+	0.75 (0.74–0.76)
Multiple DXA	Frost [45]	P	Community mailings to all aged >60 yrs without osteoporosis in one Australian community	0/750	7.1	0/9.3	60+	0.75 (0.74–0.76)
Multiple DXA	Berry [44]	P	Surviving participants in Framingham cohort without baseline fractures scanned > once from 1987 to 1999	492/310	9.6	25.3/14.1	70+	0.72 (0.66–0.79)
pDEXA	Miller [10]	P	Randomly chosen names of women at 4236 GP-offices	51,914/0	1	6.4/1.6	50+	0.69 (NR)
pDXA	Miller [10]	P	Randomly chosen names of women at 4236 GP-offices	10,836/0	1	6.4/1.6	50+	0.7 (NR)
QCT	Black [35]	P	Community mailings from six medical centers	0/3347	5.5	4.9/0.2	65+	0.86 (NR)
QUS	Miller [10]	P	Randomly chosen names of women at 4236 GP-offices	7562/0	1	6.4/1.6	50+	0.67
QUS	Dargent-Molina [26]	P	Randomly selected from voting or health care registries in France (EPIDOS)	5910/0	4	NR/3.9	75+	NR
QUS	Bauer [25]	P	Community mailings from six medical centers	0/5607	4.2	4.9/5.0	65+	0.68 (NR)
QUS	Durosier [27]	P	Pooled EPIDOS and SEMOF cohorts	12,064/0	3.2	NR/2.5	70+	0.76 (0.74–0.79)
QUS	Guessous [24]	P	Randomly selected from resident register in 9 Swiss cantons (SEMOF)	6174/0	2.8	NR/5.1	70–85	NR
QUS	Glüer [21]	CS	Recruitment from registers and primary care centers	1265/0	NA	NR/16	55–79	0.66–0.67 (NR)
QUS	Huopio [29]	P	Random subgroup of OST-PRE, recruited by mailed invitation to all female inhabitants aged 47–56 yrs in one Finnish community	422/0	2.6	NR/7.6	47–56	0.68 (NR)
QUS	Stewart [28]	P	Subgroup of random sample of women in one Scottish community	775/0	9.7	NR/10.0	45–54	0.64 (0.63–0.66)
QUS	Hollaender [30]	P	Randomly selected from resident register in one city	432	3.4	NR/5.6	60–80	0.73 (0.63–0.82)
RA	Friis-Holmberg [34]	P	Randomly selected from resident register in 13 municipalities	7552/0	4.3	12.3/4.1	18–95	0.71 (0.69–0.74)
RA	Friis-Holmberg [34]	P	Randomly selected from resident register in 13 municipalities	0/5206	4.3	3.7/1.6	18–95	0.64 (0.58–0.70)
TBS	Boutroy [41]	P	Random sample from health insurance records in Rhone	560/0	8	23.9/20.1	postmenopausal	0.60 (0.62–0.74)
TBS	Briot [40]	P	Random samples from resident or health insurance registers or from GP lists	2409/0	6	27.5/4.6	50+	0.62 (NR)
TBS	Iki [42]	P	Age-stratified random sample from seven areas of Japan	665/0	4.7	40.6/13.8	50+	0.68 (0.62–0.74)

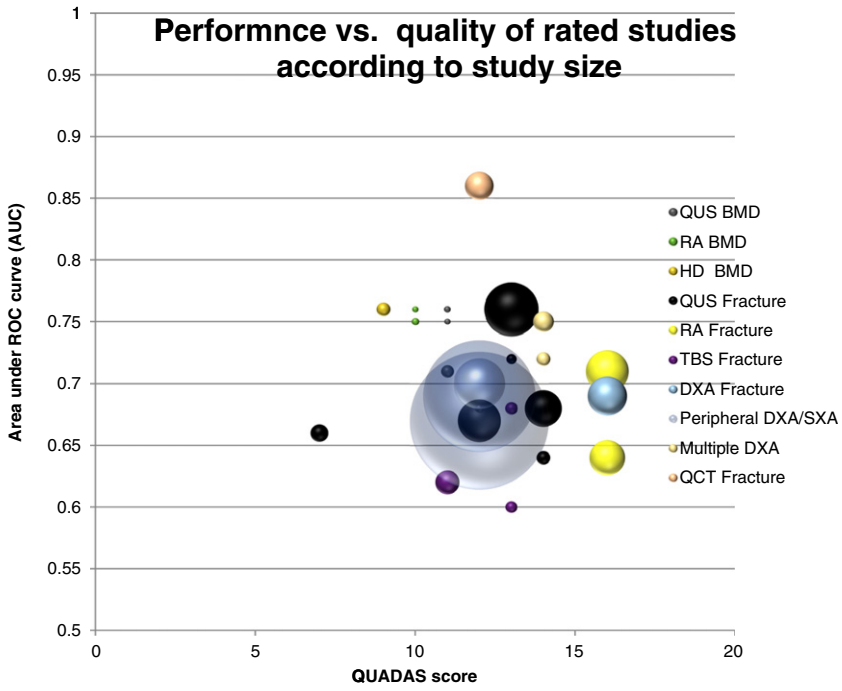

Fig. 4. Performance of diagnostic devices in each population-based study, measured by area under the curve (AUC) of a receiver-operating characteristic, by quality as measured by QUADAS II score. The size of the bubbles reflects the relative size of the study whilst different colors relates to different diagnostic devices applied.

of TBS to BMD alone and with only one of these studies [40] finding an additive effect of TBS to known BMD values.

3.14. Nontraditional DXA results

3.14.1. Prognostic use: nontraditional or derived DXA results for the prediction of future fractures

The use of alternative cutoffs for fracture risk evaluation and the use of repeat BMD measurements have been examined in a few studies. Wu et al. [43] found that mathematically established lower limit of normal by the use of NHANES data resulted in a more consistent classification of participants in osteoporotic/non-osteoporotic groups and provided better prediction of future fractures than the WHO definition of osteoporosis. Berry et al. [44] examined the additive value of change of BMD between two or more DXA evaluations compared to baseline BMD alone for prediction of fractures. AUC for prediction of future fractures was 0.72 (95% c.i. 0.66–0.79) for baseline BMD. Information about yearly change in BMD did not change the predictive value. Frost et al. [45] proposed a model to calculate the optimal time of repeat BMD testing, dependent upon sex, age and T-score at baseline. However, the model was not tested in an independent cohort. Leslie et al. [46] used T-score offset between spine and hip as a separate predictor of major osteoporotic fractures in the BMD-corrected FRAX algorithm, using observed fracture outcomes in the CaMOS cohort. Here, AUC for prediction of future fractures was 0.69 (0.67–0.71) for FRAX alone and 0.693 (0.671–0.714) for the T-score offset corrected values. No statistical difference was found, but risk of fractures could be reclassified for 5.5% of participants.

4. Discussion

During a systematic review of population-based studies that focused on technical methods that could substitute for DXA, we found that only QUS appears capable of perhaps replacing DXA as standalone examination in the future whilst radiographic absorptiometry could provide important information in areas with scarcity of DXA. QUS may be of added value even after DXA has been performed. However, evaluation of proposed cutoff-values from population-based studies in separate population-based cohorts is still lacking for most examination devices.

Using a structured literature search of OVID and Pubmed, 23 population-based studies examining other technical devices for osteoporosis using a properly documented methodology to either indicate BMD by DXA or predict fracture risk were identified. Overall, the methodological quality of prospective studies aiming at fracture prediction seemed better than cross-sectional studies with DXA measurements as outcome. Differences in follow-up and included participants were potential explanatory variables for the wide range in fractures recorded during follow-up of prospective studies.

The ability of QUS to indicate BMD by DXA and predict fractures was found to be modest. Most prospective studies found a BMD-independent predictive ability of QUS for fractures. None of the studies tested their models in separate population-based samples. These findings are in line with a recent meta-analysis of QUS aimed at exploring the feasibility of QUS as prescreening tool before DXA in postmenopausal women [47]. A consistent ability of QUS to indicate DXA defined osteoporosis was found; however with a wide range in both percentage of DXA saved and misclassification rates. Two other meta-analyses of prospective studies used gradient of risk as outcome, also showing some heterogeneity between studies, but with overall significant predictive ability for fractures even after BMD-adjustment, independent of sex [15,48].

Based on the literature review, QUS could have the possibility to supplant DXA as standalone examination due to lower cost and portability of the technology. In areas with scarcity of DXA machines, QUS could be used to reduce the number of needed DXA examinations for fracture prediction. Due to the possibly BMD-independent effect of fracture prediction found in most studies included in this review, QUS could also serve for further risk stratification in areas with ample DXA coverage, so that patients with intermediate risk of future fractures can be further subdivided in low- and high-risk individuals. However, the lack of universal calibration between scanners and lack of prospective outcome

studies where patients have been diagnosed and treated on the basis of QUS still hamper its use.

Radiographic absorptiometry in phalanges was found to have a modest ability to discriminate between osteoporotic and non-osteoporotic individuals whilst the ability of heel examinations to predict future fractures was relatively minor. Predictive algorithms have been proposed but they have not yet been validated in separate cohorts.

In areas with scarcity of DXA machines, RA might be used to reduce the number of needed DXA examinations for fracture prediction, as the equipment seems to identify a risk not captured by risk factors alone [34]. Widespread usage would be promoted if robust cross-calibration procedures as seen for DXA [8] also were available for RA.

In population-based studies, we found TBS to have an equally discriminatory value, but no additive value, for fracture discrimination as compared to BMD alone. This is in concordance with a recent review, where fracture discrimination odds-ratios by TBS were significantly attenuated after adjustment for confounders [16]. Opposing this view are important findings from the largest meta-analysis of trabecular bone score results to date, pooling individual TBS data from 14 prospective cohorts [49]. Previously, TBS data was not available from most of these cohorts. TBS adjustment factors for FRAX (including BMD) probabilities was derived from the non-population-based Manitoba bone density program and applied to pooled data from the 14 cohorts (of which the Manitoba program supplied 59% of individual data). Expressing the prognostic impact of TBS on fracture risk in hazard ratio per 1 SD change, TBS remained a significant predictor despite adjustment by FRAX (including BMD). Overall, on the finding of this meta-analysis, TBS might have a smaller impact on the diagnostic gap in osteoporosis in the future, even though the issue of relative inaccessibility of DXA equipment also remains a problem for TBS. The information derived by QCT has a potential impact on fracture prediction from a mechanical understanding, compared to the integrated measure of BMD acquired on a DXA machine. One population-based study found a good predictive ability of QCT for further fractures. However, the evaluation of QCT is time-consuming, the equipment remains expensive and radiation doses are several magnitudes larger than that of DXA. As of now, available evidence still does not support a potential for improvement of fracture prediction by the use of QCT due to the shortcomings of the methodology.

No population-based studies were found examining the ability of HR-pQCT to discriminate between persons at high or low risk of osteoporosis or fracture. Thus, whilst promising technologically due to its ability to measure bone microarchitecture, prospective studies clarifying its ability to indicate osteoporosis by DXA and/or predict fractures in a population-based setting are still needed.

A single study indicates that physical measures can indicate underlying osteoporosis with moderate precision and that it may also serve as a predictor of future fractures. Such measures could be incorporated into fracture risk algorithms in the future. However, the proposed physical measure has not been tested in a separate, prospective cohort.

There is insufficient evidence in population-based studies to alter the currently established DXA-based definition of osteoporosis or to include change in BMD or T-score offset as a separate variable.

Change in BMD between consecutive BMD measurements have not been shown to improve the predictive value of BMD for future fractures; though the study by Frost et al. [45] suggests that timing of follow-up BMD itself could be individualized a formal test of repeated assessments in a prospective cohort is still lacking.

We included only research papers in English and Nordic languages. Furthermore, we restricted the literature review only to include population-based studies, hereby excluding a lot of well-performed studies of great interest in specific disease contexts or subgroups. However, measures of diagnostic accuracy are known to vary with disease prevalence [50], so conclusions drawn in one setting (non-population-based) cannot readily be transferred to another setting (population-based) [51]. Evaluating accuracy of different technical methods for screening purposes is therefore only possible using studies from population-based settings.

A further weakness of our study was the restriction to research papers reporting discriminatory ability by a combination of sensitivity and specificity or alternatively AUC of Receiver Operating Characteristic curves values as standard, thereby leaving out important studies using diagnostic odds-ratios, hazard ratios or gradient of risk. This was done, however, on the basis of previously published statistical recommendations [52]. As discussed above, this had a noticeably impact on our findings for TBS compared to a recently published review [49]. However, the conclusion this meta-analysis based the majority of its risk prediction on an non-population-based study, which can be problematic [51], and the authors themselves speculate that the proportion of patients that would be reclassified by the adjustment of FRAX through TBS would be minor [49].

A final weakness is related to the use of QUADAS 2 criteria for the evaluation of whether included studies were population-based or not. Thus, studies that used health care registries or residential phone numbers as basis for random sampling were included in the literature review, even though, from a stringent viewpoint, this rating could be seen as subjective and questionable. However, we did not wish to exclude well-performed studies where the sampling base was as wide as possible – but restricted due to missing universal health care or other means of identifying each and every member of the nation's society.

Strength of our study is that it is a complete and extensive review of the existing literature of population-based studies using sensitivity and specificity or AUC as outcome measures. Furthermore, we rated the literature according to the QUADAS 2 criteria as recommended by the Cochrane Handbook of Systematic reviews.

In conclusion, this extensive literature review shows that many other technical devices have been tested in a population-based setting using properly documented methodology for prediction of BMD or fracture risk. Overall, only modest ability to indicate osteoporosis in terms of a low BMD measured by DXA was found in included studies. By contrast, prospective prediction of fractures, though modest, is generally in the same range as DXA. This indicates that such techniques have potential value in identifying individuals at high risk of fracture and to the extent that they provide non-overlapping risk information with DXA, they could be performed along with DXA to further refine the risk estimate. However, this may be premature as most of the included population-based studies were hypothesis-generating in their design with limited or no use of evaluation in external cohorts.

Existing population-based studies do not support that QCT, HR-pQCT and TBS may have incremental fracture risk predictive value when BMD by DXA is already known, whereas QUS may be of added value once formal evaluation has been completed. RA could serve as mobile-prescreening before DXA due to its additive effect on risk factor-based risk-estimation, but this question needs further exploration.

Still, the effect of different screening tools for the prediction of DXA measurements and fractures with post-hoc optimized cut-off values would need to be tested in separate cohorts to prevent overoptimistic accuracy results [52,53].

In the end of the day, the most important question, whether to perform screening for osteoporosis on a population-based level or not, remains unanswered. A large number of societies, including the US Preventive Services Task-Force (USPSTF) [54] and the national Osteoporosis Foundation [55], endorse population-based screening of women, and some also of men [56], above defined values of chronological age, in addition to other high-risk groups. However, recommendations are not made on the basis of studies showing a reduction in fracture incidence or fracture-related morbidity or mortality due to screening interventions, but on extrapolation of results from clinical efficacy studies of antiresorptive treatment [57]. On the basis of the same trials, the NOF has found pharmacological primary- or secondary prevention of osteoporotic fractures to be cost-effective at a 10 year risk of hip fracture above 3%. Others, such as the German Dachverband Osteologie,

recommend the use of other cutoff-values (30% 10 year risk of major osteoporotic fracture) [58]. However, the debate is still on-going, and currently the USPSTF are updating their systematic review on the issue [59].

Future research should include randomized controlled trials using other technical equipment than DXA for fracture prediction with evaluation of pre-defined cutoff-values for the chosen diagnostic approach.

When enough evidence has been collected to ascertain the most rational way of fracture prediction, screening studies for osteoporosis are needed, in order to remove the uncertainty described above about whether or not to screen for osteoporosis in a population-based setting. In times of budgetary constraints, these studies should also include cost-efficacy studies.

Disclosures

MPH reports personal fees from Lilly (past, within 36 months) and personal fees from Sanofi (past, within 36 months), outside the submitted work, MPH is a full time employee of Boehringer-Ingelheim, outside the submitted work.

KHR has no disclosures.

HAP reports personal fees from Amgen, personal fees from MSD, grants and personal fees from Eli Lilly, personal fees from Shire, outside the submitted work;

KB reports investigator fees paid to his department by Amgen (past, within 36 months), investigator fees paid to his department by Merck (past, within 36 months), investigator fees paid to his department by Novartis (past, within 36 month) and investigator fees paid to his department by NPS (past, within 36 months), outside the submitted work.

BA reports research grants from Novartis (current), personal fees from Nycomed/Takeda (past, within 36 months), personal fees from Merck (past, within 36 months), personal fees from Amgen (past, within 36 months), grants from UCB (current), outside the submitted work.

Acknowledgments

Authors roles:

Study design: MPH and KHR. Data collection: MPH. Data analysis: MPH and KHR Data interpretation: MPH, KHR, APH, KB and BA. Drafting manuscript: MPH and KHR. Revising manuscript content: MPH, KHR, APH, KB and BA. Approving final version of manuscript: MPH, KHR, APH, KB and BA. MPH takes responsibility for the integrity of the data analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bone.2016.08.011.

References

[1] Assessment of fracture risk and its application to screening for postmenopausal osteoporosis, Report of a WHO Study Group. World Health Organization Technical Report Series, 843, 1994, pp. 1–129.
[2] J.A. Kanis, F. Borgstrom, J. Compston, K. Dreinhofer, E. Nolte, L. Jonsson, W.F. Lems, E.V. McCloskey, R. Rizzoli, J. Stenmark, SCOPE: a scorecard for osteoporosis in Europe, Arch. Osteoporos. 8 (2013) 144.
[3] P. Vestergaard, L. Rejnmark, L. Mosekilde, Osteoporosis is markedly underdiagnosed: a nationwide study from Denmark, Osteoporos. Int. 16 (2005) 134–141.
[4] K.H. Rubin, B. Abrahamsen, A.P. Hermann, M. Bech, J. Gram, K. Brixen, Prevalence of risk factors for fractures and use of DXA scanning in Danish women. A regional population-based study, Osteoporos. Int. 22 (2011) 1401–1409.
[5] M.P. Hoiberg, K.H. Rubin, J. Gram, A.P. Hermann, K. Brixen, G. Haugeberg, Risk factors for osteoporosis and factors related to the use of DXA in Norway, Arch. Osteoporos. 10 (2015) 220.
[6] P.D. Miller, Underdiagnosis and undertreatment of osteoporosis: the battle to be won, J. Clin. Endocrinol. Metab. 101 (2016) 852–859.
[7] D.H. Solomon, A.R. Patrick, J. Schousboe, E. Losina, The potential economic benefits of improved postfracture care: a cost-effectiveness analysis of a fracture liaison service in the US health-care system, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 29 (2014) 1667–1674.
[8] J.T. Schousboe, J.A. Shepherd, J.P. Bilezikian, S. Baim, Executive summary of the 2013 International Society for Clinical Densitometry Position Development Conference on bone densitometry, J. Clin. Densitom. 16 (2013) 455–466.
[9] D. Marshall, O. Johnell, H. Wedel, Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures, BMJ 312 (1996) 1254–1259.
[10] P.D. Miller, E.S. Siris, E. Barrett-Connor, K.G. Faulkner, L.E. Wehren, T.A. Abbott, Y.T. Chen, M.L. Berger, A.C. Santora, L.M. Sherwood, Prediction of fracture risk in postmenopausal white women with peripheral bone densitometry: evidence from the National Osteoporosis Risk Assessment, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 17 (2002) 2222–2230.
[11] E. Seeman, P.D. Delmas, Bone quality—the material and structural basis of bone strength and fragility, N. Engl. J. Med. 354 (2006) 2250–2261.
[12] J.P. van den Bergh, T.A. van Geel, W.F. Lems, P.P. Geusens, Assessment of individual fracture risk: FRAX and beyond, Curr. Osteoporos. Rep. 8 (2010) 131–137.
[13] K.H. Rubin, T. Friis-Holmberg, A.P. Hermann, B. Abrahamsen, K. Brixen, Risk assessment tools to identify women with increased risk of osteoporotic fracture: complexity or simplicity? A systematic review, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 28 (2013) 1701–1717.
[14] F. Johannesdottir, T. Turmezei, K.E. Poole, Cortical bone assessed with clinical computed tomography at the proximal femur, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 29 (2014) 771–783.
[15] E.V. McCloskey, J.A. Kanis, A. Oden, et al., Predictive ability of heel quantitative ultrasound for incident fractures: an individual-level meta-analysis, Osteoporos. Int. 26 (2015) 1979–1987.
[16] B.C. Silva, W.D. Leslie, H. Resch, O. Lamy, O. Lesnyak, N. Binkley, E.V. McCloskey, J.A. Kanis, J.P. Bilezikian, Trabecular bone score: a noninvasive analytical method based upon the DXA image, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 29 (2014) 518–530.
[17] K.G. Faulkner, L.A. Roberts, M.R. McClung, Discrepancies in normative data between Lunar and Hologic DXA systems, Osteoporos. Int. 6 (1996) 432–436.
[18] L. Shamseer, D. Moher, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, L.A. Stewart, Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation, BMJ 349 (2015) g7647.
[19] P.F. Whiting, A.W. Rutjes, M.E. Westwood, S. Mallett, J.J. Deeks, J.B. Reitsma, M.M. Leeflang, J.A. Sterne, P.M. Bossuyt, QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies, Ann. Intern. Med. 155 (2011) 529–536.
[20] J. Higgins, S. Green, Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011.
[21] C.C. Gluer, R. Eastell, D.M. Reid, et al., Association of five quantitative ultrasound devices and bone densitometry with osteoporotic vertebral fractures in a population-based sample: the OPUS Study, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 19 (2004) 782–793.
[22] Y. Ikeda, M. Iki, A. Morita, H. Aihara, S. Kagamimori, Y. Kagawa, T. Matsuzaki, H. Yoneshima, F. Marumo, Age-specific values and cutoff levels for the diagnosis of osteoporosis in quantitative ultrasound measurements at the calcaneus with SAHARA in healthy Japanese women: Japanese population-based osteoporosis (JPOS) study, Calcif. Tissue Int. 71 (2002) 1–9.
[23] S.L. Gudmundsdottir, O.S. Indridason, L. Franzson, G. Sigurdsson, Age-related decline in bone mass measured by dual-energy X-ray absorptiometry and quantitative ultrasound in a population-based sample of both sexes: identification of useful ultrasound thresholds for osteoporosis screening, J. Clin. Densitom. 8 (2005) 80–86.
[24] I. Guessous, J. Cornuz, C. Ruffieux, P. Burckhardt, M.A. Krieg, Osteoporotic fracture risk in elderly women: estimation with quantitative heel US and clinical risk factors, Radiology 248 (2008) 179–184.
[25] D.C. Bauer, S.K. Ewing, J.A. Cauley, K.E. Ensrud, S.R. Cummings, E.S. Orwoll, Quantitative ultrasound predicts hip and non-spine fracture in men: the MrOS study, Osteoporos. Int. 18 (2007) 771–777.
[26] P. Dargent-Molina, S. Piault, G. Breart, A comparison of different screening strategies to identify elderly women at high risk of hip fracture: results from the EPIDOS prospective study, Osteoporos. Int. 14 (2003) 969–977.
[27] C. Durosier, D. Hans, M.A. Krieg, C. Ruffieux, J. Cornuz, P.J. Meunier, A.M. Schott, Combining clinical factors and quantitative ultrasound improves the detection of women both at low and high risk for hip fracture, Osteoporos. Int. 18 (2007) 1651–1659.
[28] A. Stewart, V. Kumar, D.M. Reid, Long-term fracture prediction by DXA and QUS: a 10-year prospective study, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 21 (2006) 413–418.
[29] J. Huopio, H. Kroger, R. Honkanen, J. Jurvelin, S. Saarikoski, E. Alhava, Calcaneal ultrasound predicts early postmenopausal fractures as well as axial BMD. A prospective study of 422 women, Osteoporos. Int. 15 (2004) 190–195.
[30] R. Hollaender, F. Hartl, M.A. Krieg, A. Tyndall, C. Geuckel, C. Buitrago-Tellez, M. Manghani, M. Kraenzlin, R. Theiler, D. Hans, Prospective evaluation of risk of vertebral fractures using quantitative ultrasound measurements and bone mineral density in a population-based sample of postmenopausal women: results of the Basel Osteoporosis Study, Ann. Rheum. Dis. 68 (2009) 391–396.
[31] N.J. Perkins, E.F. Schisterman, The Youden Index and the optimal cut-point corrected for measurement error, Biom. J. 47 (2005) 428–441.
[32] S. Lekamwasam, L. Wijayaratne, M. Rodrigo, U. Hewage, Prevalence of osteoporosis among postmenopausal women in Sri Lanka: a cross-sectional community study, APLAR J. Rheumatol. 10 (2007) 234–238.
[33] S.J. Hansen, M.M. Nielsen, J. Ryg, K. Wraae, M. Andersen, K. Brixen, Radiographic absorptiometry as a screening tool in male osteoporosis: results from the Odense Androgen Study, Acta Radiol. 50 (2009) 658–663 (Stockholm, Sweden: 1987).

[34] T. Friis-Holmberg, K.H. Rubin, K. Brixen, J.S. Tolstrup, M. Bech, Fracture risk prediction using phalangeal bone mineral density or FRAX®?-a Danish cohort study on men and women, J. Clin. Densitom. 17 (2014) 7–15.
[35] D.M. Black, M.L. Bouxsein, L.M. Marshall, S.R. Cummings, T.F. Lang, J.A. Cauley, K.E. Ensrud, C.M. Nielson, E.S. Orwoll, Proximal femoral structure and the prediction of hip fracture in men: a large prospective study using QCT, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 23 (2008) 1326–1333.
[36] A.W. Popp, H. Buffat, U. Eberli, K. Lippuner, M. Ernst, R.G. Richards, V.A. Stadelmann, M. Windolf, Microstructural parameters of bone evaluated using HR-pQCT correlate with the DXA-derived cortical index and the trabecular bone score in a cohort of randomly selected premenopausal women, PLoS One 9 (2014), e88946.
[37] M. Karkkainen, T. Rikkonen, H. Kroger, J. Sirola, M. Tuppurainen, K. Salovaara, J. Arokoski, J. Jurvelin, R. Honkanen, E. Alhava, Physical tests for patient selection for bone mineral density measurements in postmenopausal women, Bone 44 (2009) 660–665.
[38] W.D. Leslie, B. Aubry-Rozier, O. Lamy, D. Hans, TBS (trabecular bone score) and diabetes-related fracture risk, J. Clin. Endocrinol. Metab. 98 (2013) 602–609.
[39] S.J. Silverberg, B.L. Clarke, M. Peacock, et al., Current issues in the presentation of asymptomatic primary hyperparathyroidism: proceedings of the Fourth International Workshop, J. Clin. Endocrinol. Metab. 99 (2014) 3580–3594.
[40] K. Briot, S. Paternotte, S. Kolta, R. Eastell, D.M. Reid, D. Felsenberg, C.C. Gluer, C. Roux, Added value of trabecular bone score to bone mineral density for prediction of osteoporotic fractures in postmenopausal women: the OPUS study, Bone 57 (2013) 232–236.
[41] S. Boutroy, D. Hans, E. Sornay-Rendu, N. Vilayphiou, R. Winzenrieth, R. Chapurlat, Trabecular bone score improves fracture risk prediction in non-osteoporotic women: the OFELY study, Osteoporos. Int. 24 (2013) 77–85.
[42] M. Iki, J. Tamaki, E. Kadowaki, Y. Sato, N. Dongmei, R. Winzenrieth, S. Kagamimori, Y. Kagawa, H. Yoneshima, Trabecular bone score (TBS) predicts vertebral fractures in Japanese women over 10 years independently of bone density and prevalent vertebral deformity: the Japanese Population-Based Osteoporosis (JPOS) cohort study, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 29 (2014) 399–407.
[43] Q. Wu, J.H. Magnus, J.C. Rice, J.J. Lefante, Does using lower limit of normal values enhance the ability of a single bone mineral density measure to predict fractures? Osteoporos. Int. 21 (2010) 1881–1888.
[44] S.D. Berry, E.J. Samelson, M.J. Pencina, R.R. McLean, L.A. Cupples, K.E. Broe, D.P. Kiel, Repeat bone mineral density screening and prediction of hip and major osteoporotic fracture, JAMA 310 (2013) 1256–1262.
[45] S.A. Frost, N.D. Nguyen, J.R. Center, J.A. Eisman, T.V. Nguyen, Timing of repeat BMD measurements: development of an absolute risk-based prognostic model, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 24 (2009) 1800–1807.
[46] W.D. Leslie, C.S. Kovacs, W.P. Olszynski, T. Towheed, S.M. Kaiser, J.C. Prior, R.G. Josse, S.A. Jamal, N. Kreiger, D. Goltzman, Spine-hip T-score difference predicts major osteoporotic fracture risk independent of FRAX®: a population-based report from CAMOS, J. Clin. Densitom. 14 (2011) 286–293.
[47] K. Thomsen, D.B. Jepsen, L. Matzen, A.P. Hermann, T. Masud, J. Ryg, Is calcaneal quantitative ultrasound useful as a prescreen stratification tool for osteoporosis? Osteoporos. Int. 26 (2015) 1459–1475.
[48] A. Moayyeri, J.E. Adams, R.A. Adler, M.A. Krieg, D. Hans, J. Compston, E.M. Lewiecki, Quantitative ultrasound of the heel and fracture risk assessment: an updated meta-analysis, Osteoporos. Int. 23 (2012) 143–153.
[49] E.V. McCloskey, A. Oden, N.C. Harvey, et al., A meta-analysis of trabecular bone score in fracture risk prediction and its relationship to FRAX, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 31 (2016) 940–948.
[50] M.M. Leeflang, A.W. Rutjes, J.B. Reitsma, L. Hooft, P.M. Bossuyt, Variation of a test's sensitivity and specificity with disease prevalence, Can. Med. Assoc. J. 185 (2013) E537–E544.
[51] P.F. Whiting, A.W. Rutjes, M.E. Westwood, S. Mallett, A systematic review classifies sources of bias and variation in diagnostic test accuracy studies, J. Clin. Epidemiol. 66 (2013) 1093–1104.
[52] D.G. Altman, Y. Vergouwe, P. Royston, K.G. Moons, Prognosis and prognostic research: validating a prognostic model, BMJ 338 (2009) b605.
[53] M.M. Leeflang, K.G. Moons, J.B. Reitsma, A.H. Zwinderman, Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions, Clin. Chem. 54 (2008) 729–737.
[54] H.D. Nelson, E.M. Haney, T. Dana, C. Bougatsos, R. Chou, Screening for osteoporosis: an update for the U.S. Preventive Services Task Force, Ann. Intern. Med. 153 (2010) 99–111.
[55] F. Cosman, S.J. de Beur, M.S. LeBoff, E.M. Lewiecki, B. Tanner, S. Randall, R. Lindsay, Clinician's guide to prevention and treatment of osteoporosis, Osteoporos. Int. 25 (2014) 2359–2381.
[56] N.B. Watts, R.A. Adler, J.P. Bilezikian, M.T. Drake, R. Eastell, E.S. Orwoll, J.S. Finkelstein, Osteoporosis in men: an Endocrine Society clinical practice guideline, J. Clin. Endocrinol. Metab. 97 (2012) 1802–1822.
[57] D. Altkorn, A.S. Cifu, Screening for osteoporosis, JAMA 313 (2015) 1467–1468.
[58] D. Osteologie, DVO Leitlinie Osteoporose, http://www.dv-osteologie.org/dvo_leitlinien/osteoporose-leitlinie-20142014 (Accessed August 9th 2016).
[59] U.P.S.T. Force, Final Research Plan for Osteoporotic Fractures: Screening, http://www.uspreventiveservicestaskforce.org/Page/Document/final-research-plan/osteoporosis-screening1 (Accessed August 9th 2016).
[60] A.F. Leite, P.T. Figueiredo, C.M. Guia, N.S. Melo, A.P. de Paula, Correlations between seven panoramic radiomorphometric indices and bone mineral density in postmenopausal women, Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod. 109 (2010) 449–456.
[61] K. Karayianni, K. Horner, A. Mitsea, et al., Accuracy in osteoporosis diagnosis of a combination of mandibular cortical width measurement on dental panoramic radiographs and a clinical risk index (OSIRIS): the OSTEODENT project, Bone 40 (2007) 223–229.
[62] A. Massie, D.M. Reid, R.W. Porter, Screening for osteoporosis: comparison between dual energy X-ray absorptiometry and broadband ultrasound attenuation in 1000 perimenopausal women, Osteoporos. Int. 3 (1993) 107–110.
[63] V. Naganathan, L. March, D. Hunter, N.A. Pocock, J. Markovey, P.N. Sambrook, Quantitative heel ultrasound as a predictor for osteoporosis, Med. J. Aust. 171 (1999) 297–300.
[64] J. Nairus, S. Ahmadi, S. Baker, D. Baran, Quantitative ultrasound: an indicator of osteoporosis in perimenopausal women, J. Clin. Densitom. 3 (2000) 141–147.
[65] M. Ayers, M. Prince, S. Ahmadi, D.T. Baran, Reconciling quantitative ultrasound of the calcaneus with X-ray-based measurements of the central skeleton, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 15 (2000) 1850–1855.
[66] K.I. Kim, I.K. Han, H. Kim, N.H. Cho, How reliable is the ultrasound densitometer for community screening to diagnose osteoporosis in spine, femur, and forearm? J. Clin. Densitom. 4 (2001) 159–165.
[67] A.W. Kung, A.Y. Ho, W.B. Sedrine, J.Y. Reginster, P.D. Ross, Comparison of a simple clinical risk index and quantitative bone ultrasound for identifying women at increased risk of osteoporosis, Osteoporos. Int. 14 (2003) 716–721.
[68] A.W. Kung, A.Y. Ho, P.D. Ross, J.Y. Reginster, Development of a clinical assessment tool in identifying Asian men with low bone mineral density and comparison of its usefulness to quantitative bone ultrasound, Osteoporos. Int. 16 (2005) 849–855.
[69] K.M. Gasser, C. Mueller, M. Zwahlen, M. Kaufmann, G. Fuchs, R. Perrelet, G. Abetel, U. Burgi, K. Lippuner, Osteoporosis case finding in the general practice: phalangeal radiographic absorptiometry with and without risk factors for osteoporosis to select postmenopausal women eligible for lumbar spine and hip densitometry, Osteoporos. Int. 16 (2005) 1353–1362.
[70] R.J. Barr, A. Adebajo, W.D. Fraser, J.P. Halsey, C. Kelsey, A. Stewart, D.M. Reid, Can peripheral DXA measurements be used to predict fractures in elderly women living in the community? Osteoporos. Int. 16 (2005) 1177–1183.
[71] T.B. Brismar, I. Janszky, L.I. Toft, Calcaneal BMD obtained by dual X-ray and laser predicts future hip fractures-a prospective study on 4 398 Swedish women, J. Osteoporos. 2010 (2010) 875647.
[72] W.D. Leslie, J.F. Tsang, P.A. Caetano, L.M. Lix, Effectiveness of bone density measurement for predicting osteoporotic fractures in clinical practice, J. Clin. Endocrinol. Metab. 92 (2007) 77–81.
[73] T. Kwok, C.C. Khoo, J. Leung, A. Kwok, L. Qin, J. Woo, P.C. Leung, Predictive values of calcaneal quantitative ultrasound and dual energy X ray absorptiometry for non-vertebral fracture in older men: results from the MrOS study (Hong Kong), Osteoporos. Int. 23 (2012) 1001–1006.
[74] W.D. Leslie, B. Aubry-Rozier, L.M. Lix, S.N. Morin, S.R. Majumdar, D. Hans, Spine bone texture assessed by trabecular bone score (TBS) predicts osteoporotic fractures in men: the Manitoba Bone Density Program, Bone 67 (2014) 10–14.
[75] D. Hans, A.L. Goertzen, M.A. Krieg, W.D. Leslie, Bone microarchitecture assessed by TBS predicts osteoporotic fractures independent of bone density: the Manitoba study, J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. 26 (2011) 2762–2769.