How to Best Predict Fragility Fractures: An Update and Systematic Review

Raviv Allon BS'C', Yahav Levy MD2*, Idit Lavi MA3, Aviv Kramer MD1,2, Menashe Barzilai MD1 and Ronit Wollstein MD1,2,4,5

1Rappaport Faculty of Medicine, Technion–Israel Institute of Technology, Haifa, Israel
Departments of 2Orthopedic Surgery and 3Community Medicine and Epidemiology, Carmel Medical Center, Haifa, Israel
4Department of Plastic Surgery, University of Pittsburgh, Pittsburgh, PA, USA
5Department of Orthopedic Surgery, New York University School of Medicine, Huntington Station, NY, USA

ABSTRACT: Because fragility fractures have an enormous impact on the practice of medicine and global health systems, effective screening is imperative. Currently, dual-energy X-ray absorptiometry (DXA), which has limited ability to predict fractures, is being used. We evaluated the current literature for a method that may constitute a better screening method to predict fragility fractures. A systematic review of the literature was conducted on computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound to evaluate screening methods to predict fragility fractures. We found that ultrasound had sufficient data on fracture prediction to perform meta-analysis; therefore, we analyzed prospective ultrasound cohort studies. Six study populations, consisting of 29,299 individuals (87,296 person-years of observation) and including 992 fractures, were analyzed. MRI was found to be sensitive and specific for osteoporosis, but its use for screening has not been sufficiently evaluated and more research is needed on cost, accessibility, technical challenges, and sensitivity and specificity. CT could predict fracture occurrence; however, it may be problematic for screening due to cost, exposure to radiation, and availability. Ultrasound was found to predict fracture occurrence with an increased risk of 1.45 (95% confidence interval 1.21–1.73) to fracture. Ultrasound has not replaced DXA as a screening tool for osteoporosis, perhaps due to operator-dependency and difficulty in standardization of testing.

KEY WORDS: computed tomography (CT), fragility fractures, magnetic resonance imaging (MRI), osteoporosis, ultrasound

Osteoporosis and its most devastating consequence, fragility fractures, have had an enormous impact on global health, as well as on economic and social issues. The impact of osteoporosis will continue to increase as people’s life expectancy increases [1]. In the United States, 7 million to 12 million people are estimated diagnosed with osteoporosis and another 29 million to 47 million present with low bone mass [2]. One of the largest health maintenance organizations in the world, which insures about 25% of the population in Israel, recently conducted an osteoporosis registry and found 118,141 insured patients with a diagnosis of osteoporosis, suggesting that overall about 500,000 people are diagnosed with osteoporosis in the Israeli population [3,4]. For obvious reasons, much effort has been made for the effective prevention of fragility fractures, including fall prevention and medical and physical treatment of osteoporosis.

Fragility fractures occur when low energy is exerted on a compromised bone. The bone may be compromised due to loss of mass or lower bone quality, specifically deterioration of the bony macro- and micro-architecture. To date, the gold standard for osteoporosis diagnosis is the measurement of bone mineral density (BMD), which measures bone quantity using dual-energy X-ray absorptiometry (DXA) [5]. The DXA system scans the body area in a rectilinear fashion and records separate low- and high-energy transmitted photon intensity values on a pixel-by-pixel basis. The program determines cross-sectional properties at various levels of the bone and calculates strength estimates at each section. The use of computer software allows more rapid and precise measurements of the bone involved [6].

The United States Preventive Services Task Force (USPSTF) recommends screening for osteoporosis with DXA for women aged 65 years and older and for younger women with increased risk for osteoporosis [6]. However, recently limitations to the use of DXA as a screening tool have been identified, prompting a search for a better method.

DXA equipment is heavy and non-portable, involves radiation exposure, and is often restricted to tertiary care hospitals or specialized clinics. Dedicated trained personnel are required for proper operation, which further restricts its accessibility to the general public. More importantly, some studies have questioned the ability of DXA to predict the occurrence of fragility fractures. Recent studies have suggested the addition of focused questionnaires, such as the fracture risk assessment tool (FRAX) and other risk scores as an adjunct to screening with DXA to increase the ability to predict fractures [7]. Other methods to
evaluate osteoporosis have also been suggested, some of which are based on trabecular architecture or bone quality to determine the probability of fragility fractures rather than bone quantity alone [8].

Due to the increasing incidence and medical, social, and economic significance of osteoporosis, screening and prevention of the condition is becoming more important. A superior screening tool should be inexpensive and easily accessible with reproducible results that could help in the prediction of fragility fractures. The goal of our review and limited meta-analysis was to evaluate the current literature to determine a method that may improve on DXA as the standard screening test, specifically in predicting fragility fractures. Specifically, we examined the literature on quantitative computed tomography (QCT), magnetic resonance imaging (MRI), and ultrasound.

PATIENTS AND METHODS

SEARCH STRATEGY

We performed a comprehensive search using computer databases of medical literature (PubMed and Google Scholar) for the years 2004–2017. We included all articles that dealt with both osteoporosis and the outcome measure of fragility fractures. Our search was based on the following search terms: bone density, fracture risk, osteoporosis, quantitative ultrasound, quantitative computed tomography (QCT), fragility fracture, DXA, MRI, and FRAX. In addition, the references of the retrieved articles were reviewed to uncover other relevant studies.

The search was limited by language (English only) and type of publication. No letters, editorials, or case reports were included. Two authors independently reviewed all studies for eligibility based on their title and abstract. The full text of the included studies were then assessed.

OUTCOME MEASURES

Studies were selected in the first stage for the outcome measures of osteoporosis (usually identified by DXA) and fragility fractures.

Because our aim was to find an optimal screening test to predict fractures, we initially reviewed articles describing the ability to predict osteoporosis. These articles were part of the review but not part of our analysis. During a second stage, the evaluators removed articles that did not include the outcome measure of fracture occurrence. These studies were then analyzed.

DATA EXTRACTION

Data was extracted by one of the investigators and reviewed by a second investigator. The methodological quality of the studies was assessed by the Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA-P) Statement.

We used the PRISMA-P Statement because of the prospective cohort studies dealing with our hypothesis. Levels of evidence and the risk of bias were assessed independently by two of the authors. Data were then extracted, including study design; demographic information regarding study size, age, gender, type of fracture; and follow-up period. The primary outcome measure was identified as the ability to predict future fractures.

Since the modalities reported to predict fracture occurrence were mostly part of studies evaluating ultrasound or QCT, our analysis was performed on these two modalities. However, we found that these studies were not homogenous with regard to location or precise method of ultrasound or CT evaluation and therefore could not be compared. Consequently, we were only able to analyze a small fraction of the studies that evaluated ultrasound. However, we did review all suitable modalities in our study to consider the effectiveness of ultrasound and QCT, in general, to predict fracture occurrence.

STATISTICAL ANALYSIS

We only evaluated prospective cohort studies that calculated the hazard ratios for fracture occurrence and also included only those studies that analyzed data as continuous variables, calculating the differences in one standard deviation. Studies that described fractures in different body areas were analysed by creating a variable that included all fracture locations. In studies in which different ultrasound methods were used for the same site but did not coincide with the same participants, hazard ratio estimates were combined into one estimate. The pooled effect across studies was calculated.

Statistical heterogeneity was quantified using two measures of heterogeneity: H and I^2. H < 1.2 suggests absence of noteworthy heterogeneity, whereas H > 1.5 suggests its presence. I^2 describes the percentage of variation across studies that are caused by heterogeneity rather than chance. Due to detection of significant heterogeneity, we used random effect models in our analysis. Fracture occurrence was analysed as any fracture rate (including all fractures) and a separate analysis was performed for the prediction of hip fractures.

RESULTS

Our original search yielded 879 studies. After filtering and removing duplicates when evaluating different methods of predicting both osteoporosis and fracture occurrence, 73 studies remained. These articles were further trimmed to 11 studies, with fracture occurrence as the main outcome measure. We found no studies evaluating the ability of MRI to predict fracture occurrence. The modalities that could be investigated for their ability to predict fracture occurrence were QCT and ultrasound. However, the disparities in technique, outcomes measured, and
patients were evaluated for fracture development, which cohort studies included in the meta-analysis. In all, 29,299 [Figure 1] [9-14].

methods of analysis prevented us from performing any meaningful analysis on QCT or for many of the studies published on ultrasound. Of the studies evaluating ultrasound, only six included hazard ratios and were similar in site and method/technique used. These studies were used for the meta-analysis [Figure 1] [9-14].

Table 1 and Table 2 describe the characteristics of the cohort studies included in the meta-analysis. In all, 29,299 patients were evaluated for fracture development, which accounted for approximately 87,296 person-years of observation. We observed 992 fractures: 186 wrist fractures, 107 hip fractures, 51 humerus fractures, 648 non-vertebral fractures, and other unspecified locations.

**HETEROGENEITY AND PUBLICATION BIAS**

The proportion of variation attributable to heterogeneity for any fractures was $I^2 = 82.6\%$ (95% confidence interval [95%CI] 63.1–91.8). $H$ was 2.4 (95%CI 1.6–3.5). For hip fractures, the proportion of variation attributable to heterogeneity was $I^2 = 81.3\%$ (95%CI 51.4–92.8). $H$ was 2.3 (95%CI 1.4–3.7).

**THE ABILITY OF ULTRASOUND TO PREDICT FRACTURE OCCURRENCE**

We found that ultrasound was a good predictor of any fracture occurrence with an increased risk of 1.45 (95%CI 1.21–1.73) to fracture. The ability to predict a hip fracture was not significant but showed a trend: 1.52 (95%CI 0.94–2.48). Figure 2 and Figure 3 are forest plots that describe the results of the meta-analysis.

**DISCUSSION**

**MAGNETIC RESONANCE IMAGING**

While MRI may be better able to identify osteoporosis and may possibly be able to detect bony changes earlier, there is not enough evidence to support its use in predicting fracture
occurrence [15,16]. Besides its high resolution and potential to become a more sensitive and specific test, MRI requires no ionizing radiation, which makes it a desirable candidate for osteoporosis diagnosis and, especially, regular screening exams. Some recent studies have reviewed the ability of MRI to diagnose osteoporosis based on the techniques used [17]. An ultrashort echo time (UTE) sequence reflects the water content of cortical bone, and more accurately the Haversian and lacunar-canalicular system pores. This suggestion has been proposed as a surrogate marker of bone porosity and has been shown to correlate with fracture toughness properties. Furthermore, bone marrow adiposity, best described by MRI sequences, has been shown to be associated with lower BMD and higher fracture prevalence. Micro-MRI (μMRI) provides architectural information for both cortical and trabecular bone properties and has been shown to discriminate between patients with and without fragility fractures compared to BMD.

One of the main obstacles in research of MRI in the detection of osteoporosis is its operating costs and limited availability. However, new developments in this field, such as a 30 seconds 2D UTE could lower the costs dramatically.

**QUANTITATIVE COMPUTED TOMOGRAPHY**

QCT allows the distinction of cortical and trabecular volumetric BMD, thus predicting the strength of the whole bone. Trabecular and mean volumetric BMD have been shown to discriminate between patients with and without hip and vertebral fractures. QCT-based finite element analysis (FEA) also has been shown to discriminate between patients with and without hip and vertebral fractures [17]. High resolution peripheral QCT (HR-pQCT) measurements at the tibia and distal radius have been associated with any type of fractures [18].

The data in our literature review support the use of QCT to evaluate bone quality and quantity. Most prospective studies are nested within a large observational study describing the epidemiology of fracture and osteoporosis in older men (MrOS) [19] and in an age gene/environment susceptibility-Reykjavik study (AGES-REYKJAVIK) [20]. The specific study populations may lead to a selection bias. The majority of these studies demonstrated the ability of femoral QCT to predict hip fractures using different methods, among them cortical measurements such as cortical volume percent [21,22], cortical volumetric BMD [23,24], and cortical thickness. The hazard ratios (HRs) for fracture occurrence for...
these measurements were 3.2 (95% CI 2.2–4.6), 1.9–3.6 (95% CI 1.26–6.9), 1.7–5.3 (95% CI 1.3–13.72), respectively.

Trabecular measurements, such as trabecular area bone mineral density (BMD) and trabecular volumetric BMD [23,25], with HR of 1.7 (95% CI 1.2–2.4) and 1.0–6.91 (95% CI 0.64–15.35), respectively, were noted. In addition to other measurements, such as minimal cross sectional area, femoral strength, and load-to-strength ratio, HR of 1.6 (95% CI 1.2–2.1), 13.1 (95% CI 3.9–43.5), and 4.0 (95% CI 2.7–6), respectively, were determined [26]. We noted a significant disparity among the studies with regard to the anatomical location being examined. The regions ranged from the femoral neck to the trochanteric region. Yang et al. [27] presented a thorough examination of different scan locations and determined an HR of 6.20 (95% CI 2.71–14.18) using trabecular volumetric BMD scan in the medial trochanter in contrast to HR of 1.73 (95% CI 0.96–3.09) in the lateral trochanter.

Other studies evaluated the ability of vertebral QCT to predict vertebral, wrist, and hip fractures using trabecular areaal BMD and volumetric BMD [28]. Tibial cortical bone mass and area were also independently associated with any fracture risk with HR of 2.07 (95% CI 1.58–2.70) and 2.05 (95% CI 1.58–2.65), respectively [29].

The main weaknesses of QCT in terms of use as a screening tool are its high cost and the possible exposure to ionizing radiation. The required effective dose is between 1.5 to 3 millisievert (mSv), although some protocols used lower voltage levels and allowed an effective dose of less than 1 mSv as opposed to the effective dose of DXA, which is between 0.009 and 0.027 mSv [30,31]. Another weakness is the complexity of the analysis required, which necessitates specialized software.

QCT has been slow to be implemented into clinical practice and requires further study. This modality remains a promising tool for the screening of fragility fractures despite its disadvantages.

**ULTRASOUND**

We found that ultrasound was deemed a good predictor of any fracture occurrence with an increased risk of 1.45 (95% CI 1.21–1.73) to fracture. Ultrasound has been evaluated as a predictor of fragility fractures with enough comparable studies to perform a meta-analysis. Our analysis of prospective cohorts using ultrasound showed that ultrasound is more predictable for fractures other than hip fractures, although it is possible that we did not incorporate enough data to detect a significant association with hip fractures (only 107 hip fractures). Furthermore, we would have expected hip fractures to be easier to predict since they tend to occur in older patients. Wrist fractures are predictive of a consequent hip fracture [32]. A meta-analysis by Marin et al. [33], which included three of our studies in its analysis, found a significant relationship between ultrasound diagnosis and hip fractures. The difference in their findings and ours may stem from the number of patients evaluated. The Marin study included significantly more hip fractures. Another meta-analysis that evaluated only prospective studies showed a significant association between a low bone density shown on ultrasound and the occurrence of fragility fractures [34].

In our study, we did not evaluate the specific ultrasound modalities and techniques since we were looking at the feasibility of ultrasound in general as a screening tool. We, therefore, did not differentiate in our analysis between the different techniques and included studies that evaluated calcaneal bone and distal radius bone as opposed to previous meta-analyses. This finding may explain some of the differences found between our review and others. The 2015 International Society for Clinical Densitometry (ISCD) conference suggested only heel ultrasound for the management of osteoporosis [17].

Our analysis, as well as that of others, supports the use of ultrasound in the prediction of and screening for fragility fractures. The advantages of using ultrasound are its low cost, portability, and lower exposure to ionizing radiation. These features are especially important when searching for a technology that can assess large, widespread populations. However, there are certain limitations to this modality, which may explain why, despite multiple studies demonstrating a good ability to predict fractures and despite much effort into standardization, this modality has been slow to supplant DXA as the main screening tool for predicting fragility fractures.

**Magnetic resonance imaging, although likely sensitive and specific, still needs to be evaluated for fracture prediction**

Magnetic resonance imaging (MRI) is a powerful tool for the screening of fragility fractures despite its disadvantages. Magnetic resonance imaging, although likely sensitive and specific, still needs to be evaluated for fracture prediction.

**The limitations of ultrasound include:**

- Different types/techniques of examinations and at different sites in the body
- Various types of equipment and technology that are not necessarily transferrable or comparable
- Need for a skilled operator to perform the screening

In an attempt to mitigate this potential source of variation, a computer-based targeting system is used so that the transducer can be applied to the same area [35]. Although this system may improve reproducibility, it reduces the portability and ease of use. It is also not clear from the literature how much training is necessary, what the learning curve is, and if experience affects the results of the measurements.

The speed of sound (SOS) measurement performed along the length of a bone provides information about the bone density as well as the micro-architecture, cortical thickness, and elasticity of the bone [36]. Olszynski and colleagues [37] published normative data regarding SOS measured at the calcaneus using the Sunlight Omnisense portable QUS device (Sunlight Medical, Rehovot, Israel). The authors did not determine whether these data should be adjusted for distinct normal populations. Their conclusions from another comparative study stated that QUS
may be helpful as an alternative to DXA when the latter is not available, or as an adjunct to DXA in other circumstances. They concluded that QUS cannot supplant DXA as the gold standard screening tool [37]. Rivas-Ruiz and co-authors [38] compared normative data for SOS in the pediatric population in different countries. Sandstrom's group [39] established normative data for 25 year old women for calcaneal SOS, broadband ultrasound attenuation, and heel stiffness index. All but one study evaluated values obtained over the calcaneus. The Olszynski team evaluated the calcaneus, the distal radius, and the proximal tibia.

Most studies supported the addition of questionnaires evaluating risk factors such as gender, body mass index, and smoking status as adjuncts to imaging. These variables seem to add to the ability to predict fracture regardless of the modality used.

Currently, the gold standard remains DXA. It is possible that the inclusion of multiple prediction modalities in an algorithm may increase our ability to capture and treat those patients at high risk for fracture at an earlier stage. We suggest the adoption of a location-specific or institution-specific algorithm for the early detection of osteoporosis. The first step should be noninvasive and include a risk questionnaire. Those patients whose score ranks them as high risk for the development of osteoporosis may be recommended for further testing, including ultrasound. Those patients with low bone density on ultrasound can receive treatment or undergo testing with DXA.

The age for screening should depend on the population. For example, in our medical system we demonstrated a high prevalence of osteoporosis, osteopenia, and fragility fractures at very early ages (40–50 years). Furthermore, the available modalities are system specific, especially with regard to MRI and ultrasound. These algorithms need to be implemented and tested in each specific medical system. The recent Organization for Economic Co-operation and Development (OECD) report, *Health at a Glance*, placed Israel at the bottom of OECD states regarding medical technologies with only 4.1 MRI units and 9.8 CT units per million. For comparison, the United States has 39 MRI units and 41 CT units per million [40]. This finding suggests that regardless of the current situation, there are not enough MRI or CT units to use as screening tools in Israel; therefore, the use of risk questionnaires, ultrasound, and DXA are more effective screening tools. However, the use of MRI or CT could be considered in the United States. Nevertheless, the Israeli health system has an advantage in that every citizen or permanent resident is registered in one of four health plans. Therefore, implementation of a screening program on a national level could be simpler and more easily comprehensive than in countries with a wider variety of insurance options and health organizations such as the United States.

**CONCLUSIONS**

The capacity to predict the occurrence of fragility fractures can profoundly affect the health of the growing aging population. Screening is consequently imperative, especially in light of the limitations of DXA. While ultrasound and QCT have the capability to predict fragility fractures, more research is needed to further standardize and systematize these tools. Additional modalities should be examined that consider cost, accessibility to the health system, technical challenges, and sensitivity and specificity of the instrument.

**Correspondence**

Dr. R. Wollstein
Dept. of Orthopedic Surgery, New York University, Huntington Station, NY 11746, USA
Phone: (1-631) 425-2322, Fax: (1-631) 425-2191
e-mail: ronitwollstein@gmail.com

**References**


---

### Capsule

**Transfusions for preterm babies**

Platelets are immune cell fragments that act like molecular band-aids to control bleeding and help blood to clot. Premature babies can have abnormally low platelet numbers (thrombocytopenia) in the days after birth and are often given platelet transfusions to help prevent infections. **Curley** and co-authors studied more than 600 babies in a randomized clinical trial to determine how low platelets have to get to warrant intervention. Only those infants with severely low platelet levels (< 25,000/mm³; normal is approximately 150,000) benefited from transfusion. By contrast, thrombocytopenic babies with somewhat higher platelet counts (> 50,000/mm³) who received more transfusions had poorer outcomes and an increased rate of death. These surprising findings suggest that not all babies with low platelet counts should receive a prophylactic transfusion, which should lead to safer management of premature babies.

*NEJM 2018: 10.1056/NEJMoa1807320*

Eltan Israeli

---

"The recipe for perpetual ignorance is: be satisfied with your opinions and content with your knowledge"—Elbert Green Hubbard, (1866–1915), American writer, publisher, artist, and philosopher

779