# UC Berkeley
## UC Berkeley Previously Published Works

**Title**

Biomechanical Computed Tomography analysis (BCT) for clinical assessment of osteoporosis

**Permalink**

https://escholarship.org/uc/item/1hr8s3n2

**Journal**

Osteoporosis International, 31(6)

**ISSN**

0937-941X

**Authors**

Keaveny, TM
Clarke, BL
Cosman, F
et al.

**Publication Date**

2020-06-01

**DOI**

10.1007/s00198-020-05384-2

Peer reviewed

**REVIEW**

# Biomechanical Computed Tomography analysis (BCT) for clinical assessment of osteoporosis

T.M. Keaveny[1] · B.L. Clarke[2] · F. Cosman[3] · E.S. Orwoll[4] · E.S. Siris[5] · S. Khosla[2] · M.L. Bouxsein[6]

## Abstract

The surgeon general of the USA defines osteoporosis as "a skeletal disorder characterized by compromised bone strength, predisposing to an increased risk of fracture." Measuring bone strength, *Biomechanical Computed Tomography* analysis (BCT), namely, finite element analysis of a patient's clinical-resolution computed tomography (CT) scan, is now available in the USA as a Medicare screening benefit for osteoporosis diagnostic testing. Helping to address under-diagnosis of osteoporosis, BCT can be applied "opportunistically" to most existing CT scans that include the spine or hip regions and were previously obtained for an unrelated medical indication. For the BCT test, no modifications are required to standard clinical CT imaging protocols. The analysis provides measurements of bone strength as well as a dual-energy X-ray absorptiometry (DXA)–equivalent bone mineral density (BMD) T-score at the hip and a volumetric BMD of trabecular bone at the spine. Based on both the bone strength and BMD measurements, a physician can identify osteoporosis and assess fracture risk (high, increased, not increased), without needing confirmation by DXA. To help introduce BCT to clinicians and health care professionals, we describe in this review the currently available clinical implementation of the test (VirtuOst), its application for managing patients, and the underlying supporting evidence; we also discuss its main limitations and how its results can be interpreted clinically. Together, this body of evidence supports BCT as an accurate and convenient diagnostic test for osteoporosis in both sexes, particularly when used opportunistically for patients already with CT.

## Mini Abstract

Biomechanical Computed Tomography analysis (BCT) uses a patient's CT scan to measure both bone strength and bone mineral density at the hip or spine. Performing at least as well as DXA for both diagnosing osteoporosis and assessing fracture risk, BCT is particularly well-suited to "opportunistic" use for the patient without a recent DXA who is undergoing or has previously undergone CT testing (including hip or spine regions) for an unrelated medical condition.

**Keywords** Biomechanical computed tomography · Bone mineral density · Bone strength · Finite element analysis · Fracture risk · Osteoporosis

---

✉ T.M. Keaveny
tonykeaveny@berkeley.edu

1 Departments of Mechanical Engineering and Bioengineering, University of California, Berkeley, CA, USA

2 Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Mayo Clinic, Rochester, MN, USA

3 Department of Medicine, Columbia University College of Physicians and Surgeons, New York, NY, USA

4 Bone and Mineral Unit, Oregon Health and Science University, Portland, OR, USA

5 Toni Stabile Osteoporosis Center, Department of Medicine, Columbia University Medical Center, New York, NY, USA

6 Orthopedic Biomechanics Laboratory, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA

## Introduction

The surgeon general of the USA defines osteoporosis as "a skeletal disorder characterized by compromised bone strength, predisposing to an increased risk of fracture" [1, 2] and the *International Osteoporosis Foundation* defines osteoporosis as "a condition where bones become thin and lose their strength" [3]. Since it has not yet been possible clinically to measure a patient's bone strength non-invasively, osteoporosis is usually diagnosed by measuring bone mineral density (BMD) using dual-energy X-ray absorptiometry (DXA). This approach is limited in two ways. First, rates of diagnostic testing by DXA are low. In particular, each year, only 9.5% of eligible Medicare women and 1.7% of men in the USA get diagnostically screened for osteoporosis by DXA [4]. That low screening rate is of concern because it hinders osteoporosis care [5–8] and is thought to contribute to the now rising incidence rate of hip fracture in the USA [9]. This under-diagnosis problem is widely recognized [4–8, 10–12] and is urgent because the size of the elderly population is continuing to increase [13].

A second limitation with current osteoporosis assessment is that the BMD measurement from DXA does not directly measure bone strength, the subject of osteoporosis. While bone strength does indeed correlate with BMD [14], a DXA-derived BMD measurement does not mechanistically capture potentially important elements of bone strength such as the bone's overall shape and three-dimensional geometry, the relative amount of cortical and trabecular bone, local variations in cortical thickness, and the internal spatial distribution of bone density. This limitation partly explains why DXA has limited sensitivity [15–20] for correctly predicting who will fracture.

Given these limitations, it is significant that a well-validated, convenient diagnostic test for osteoporosis that non-invasively assesses bone strength is now available clinically in the USA as a reimbursed Medicare screening benefit for osteoporosis. Formally referred to by the American Medical Association as "Biomechanical Computed Tomography" analysis (BCT), the test comprises a finite element analysis of bone strength using as input a clinical resolution CT scan [21]; it also includes CT-based measurements of BMD and DXA-equivalent hip BMD T-scores. First reported in 1991 [22] and used since by multiple groups in research settings—extensive reviews are provided elsewhere [23–26]—the finite element analysis component of BCT represents a "virtual stress test" that provides a functional non-invasive assessment of the breaking strength of the patient's hip (proximal femur) or spine (vertebral body). Currently, the only clinically available, FDA-cleared implementation of BCT in the USA is by the VirtuOst® software (O.N. Diagnostics, Berkeley, CA), a regulated class-II medical device that is the focus of this report.

Importantly for patient convenience, the VirtuOst implementation of BCT can utilize most hip- or spine-containing CT scans taken previously for any medical indication, without requiring any change to how those CT scans are originally acquired. Used in this way—so-called opportunistic use—the patient does not need to undergo any extra imaging for the BCT test and there is zero radiation exposure associated with the BCT test per se. Millions of patients in the osteoporosis demographic are scanned with CT covering the hip or spine each year. For example, in the US Medicare population in 2018, 6.8 million reimbursed CT exams of the abdomen or pelvis were performed and 2.6 million DXA exams [27]. Assuming that 40% of those DXA exams were taken for diagnostic screening purposes [4], these data imply that over sixfold more hip-containing CT exams were performed in 2018 than diagnostic DXA exams. Thus, opportunistic BCT could have appreciable clinical impact if widely used for diagnostic screening purposes in the older CT patient population. This review focuses primarily on this opportunistic use of BCT.

To help introduce the BCT test to clinicians and health care professionals, we review here the VirtuOst BCT test, describe how it can be used to manage patients, and suggest how its results are best interpreted in a clinical setting. We also review the available supporting evidence, with additional detail provided in a series of appendices (see Supplementary Materials). Lastly, we also discuss some key clinical issues that arise when using BCT to manage patients for osteoporosis, including the test's main limitations. Unlike other technical reviews of the finite element analysis component of BCT [23–26], unless noted otherwise, all data reviewed here relate to the VirtuOst implementation of BCT or its earlier versions, with a focus on its use in a clinical setting.

## BCT as a clinical test

### What is the BCT test?

Utilizing the information in the patient's CT scan as input, BCT performs a virtual stress test to compute a measurement of bone strength, which is the force (in units of newtons) required to virtually break or fracture the patient's hip or spine in a standardized loading configuration. The virtual stress test combines advanced medical image processing, principles of bone biomechanics, and the well-established engineering structural analysis technique of non-linear finite element analysis to simulate what is thought to characterize a typical fracturing event: a sideways fall for hip fracture and a compressive overload for vertebral fracture (Fig. 1). The VirtuOst BCT test also provides measurements of BMD that are statistically equivalent to those provided by either DXA at the hip or
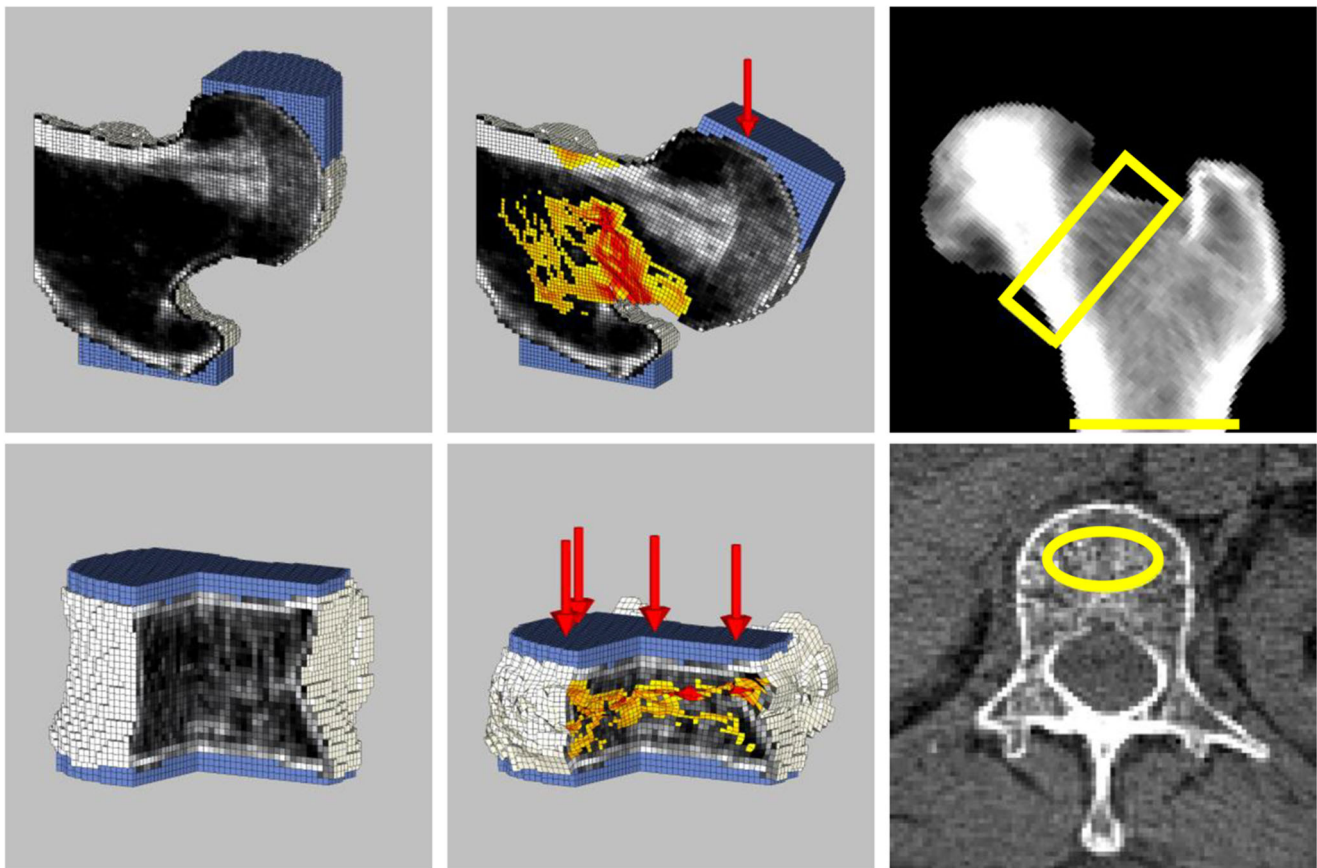
**Fig. 1** The BCT measurements for the hip (top) and spine (bottom). The finite element models (sectioned to show internal detail) depicting: (left) the spatial distribution of BMD (in grayscale) before virtual loading, and (center) the deformed shape and failed tissue (in colors) after virtual loading; deformed shape is amplified for clarity. Right: the regions of interest (in yellow) for the BMD measurements

quantitative CT at the spine (Fig. 1). Overall, the following measurements are provided at the hip or spine or both, depending on the type of CT scan used as input:

- Hip measurements: proximal femoral strength for a standardized sideways fall; DXA-equivalent BMD T- and Z-scores for the femoral neck and total hip regions (can use the NHANES or other reference populations); plots of proximal femoral strength and femoral neck BMD T-score versus age with reference population means and standard deviations. Since the BMD T-score from BCT is statistically equivalent to that from DXA, it can be used with FRAX® [28] or other risk calculators.
- Spine measurements: vertebral strength for a compressive overload; vertebral trabecular volumetric BMD, with Z-scores; plots of vertebral strength and BMD versus age with reference population means and standard deviations. Note, as per clinical guidelines, T-scores are not used with vertebral trabecular BMD [29, 30]. As discussed below,

the BCT volumetric BMD spinal measurement may have some advantages over a DXA spine BMD measurement because the volumetric measurement is minimally influenced by typical degenerative changes in the posterior elements, on the vertebral surfaces, or in the adjacent vasculature, any of which can compromise the accuracy of the DXA measurement.
- Classification of fracture risk: based on these hip and spine measurements and using clinically validated cut-points, various classifications are provided to arrive at an overall fracture risk classification for the patient (high, increased, or not increased).

## Details of the (VirtuOst) BCT test

Described in detail elsewhere [31, 32], starting with a hip- or spine-containing routine clinical CT scan, the target bone in

the patient's CT scan is first identified—a proximal femur (nominally the left femur) or a single vertebral body (preferably L1, or any one level within T12 to L3). The bone is then isolated from the surrounding tissues and organs (posterior elements are virtually removed for the spine) using advanced image processing. Unlike with spinal DXA for which four vertebral levels are typically assessed, analysis of just one [33, 34] or two [35] vertebral levels (between T12 and L3) by BCT has been shown to be effective for spinal fracture prediction; similarly, analysis of just one femur is adequate for hip fracture prediction [16, 20, 35]. To provide a patient-specific calibration and therefore enable a diagnostic-quality measurement of BMD, the CT scan is calibrated using either an external calibration phantom (typical in research studies) or a phantomless approach (typical in clinical practice) using internal tissues as references (details below). Next, the isolated bone is registered into a standardized coordinate system either by mapping the patient's proximal femur onto a reference femur that is already in a standardized orientation for virtual loading or by ensuring the vertebral endplates are horizontally oriented. For monitoring changes over time, at each time point, the isolated bones can be virtually registered to each other to optimize precision for measuring temporal changes. The isolated, calibrated, and registered bone is then converted into a finite element model comprised of 1.0-mm-sized cube-shaped, eight-node brick elements. Models for the proximal femur, for example, typically have 100,000–200,000 finite elements.

For each finite element in the model, element-specific biomechanical elastic and strength material properties are generated directly from the calibrated CT scan data, based on empirical relations obtained from biomechanical testing of human cadaveric bone specimens [36–39]. When interpreting results from a BCT test, it is noteworthy that those relations reflect the biomechanical behavior of typical human bone. These relations therefore do not capture any molecular level defects that might occur in some individuals, for example, patients with collagen or mineral deficiencies that can appreciably alter the mechanical properties independent of the BMD. As a result, for these patients, the assumed relation in the BCT model between mechanical properties and BMD at the tissue level could differ from the true relation, which could introduce an error in their BCT strength measurement.

After constructing the finite element model, loading conditions are applied. The VirtuOst implementation of BCT simulates a uniform compressive over-loading of the spine or a sideways fall of the hip, both applied via a thin layer of plastic over the bone surface to mimic laboratory biomechanical testing conditions (Fig. 1). Finally, a computationally non-linear finite element analysis is performed, solving hundreds of thousands of simultaneous equations multiple times per simulation, and the resulting non-linear force-deformation curve is processed to provide the output measurement of the whole-bone breaking strength. Ongoing research is addressing the potential clinical utility of accounting for different types of loading configurations in this process, including applying multiple different forces and simulating dynamic impact [40–42]. Thus, as the BCT technology continues to evolve, if proven beneficial, some aspects of the virtual stress testing may change.

When BCT is used opportunistically (e.g., using CT scans not acquired specifically for bone assessment), unique technical challenges arise related to ensuring the following: (1) adequate quality of the image and proper quantitative calibration of the scan and (2) consistency of BCT results across patients, especially when different CT scanners and acquisition settings are used, as is typical for opportunistic use. For example, too much image noise, including metal artifacts, can compromise the calibration and overall analysis; sharp kernels (e.g., the "bone" kernel) or unusual reconstruction filters can distort the underlying grayscale data [43]; and low voltage settings ($\leq 80$ kVp) can lead to excessive noise. All these factors would disqualify an analysis if the artifacts are manifested in the calibration reference tissues or the bone of interest; typically, for example, scans are not analyzed if there is any metal in the transverse plane of the bone of interest, e.g., a hip prosthesis in either proximal femur or a posterior fixation rod that spans T12 to L3. That said, most current clinical CT scans do not exhibit these characteristics and therefore 85–95% of scans can be processed. Intravenous contrast is not a problem for BCT at the hip but can compromise a spine analysis [44], and therefore, BCT is not typically recommended for a spinal scan acquired with intravenous contrast [45]. Excessive degenerative changes do not invalidate an analysis but can require additional image processing and more nuanced clinical interpretation. Hip scans that do not extend sufficiently toward the lesser trochanter cannot be used for a bone strength analysis, although the femoral neck BMD can be measured from slightly shorter scans.

Since BCT as a clinical test is new, widespread standards and practice guidelines do not yet exist. One immediate challenge is to ensure that the highly technical BCT analysis is properly executed and that results remain consistent across software updates, over time, and when obtained by different technicians and on different CT scanners. The FDA-regulated nature of the VirtuOst software and the associated software engineering controls ensure that results remain consistent across software updates and over time; the software algorithms also account for different CT scanner characteristics via manufacturer- and acquisition-specific adjustments in the calibration process. In addition, the VirtuOst test is currently only available via a centralized laboratory service (O.N. Diagnostics, Berkeley, CA). For that service, scans are sent to the laboratory for BCT analysis, where uniquely trained technicians perform the analysis under strict controls. This overall approach helps ensure that all VirtuOst-based BCT

**Table 1** Interventional thresholds for bone strength [35] (in newtons, N) and BMD (in dimensionless DXA-equivalent T-score units for the hip or mg/cm$^3$ for the spine). A patient is considered at *high risk of fracture* if any measurement falls within the italicized entries. The bone strength thresholds were developed based on a statistical correspondence between BMD and bone strength, measured in hundreds of patients in cohorts independent of those used for their prospective validation. The BMD thresholds are based on established guidelines. See Appendix C for further details and the related validation studies

| Classification | Women | | Men | |
|---|---|---|---|---|
| | Hip | Spine | Hip | Spine |
| Fragile bone strength | *≤ 3000 N* | *≤ 4500 N* | *≤ 3500 N* | *≤ 6500 N* |
| Low bone strength | > 3000 to < 4000 | > 4500 to < 6000 | > 3500 to < 5000 | > 6500 to < 8500 |
| Normal bone strength | ≥ 4000 | ≥ 6000 | ≥ 5000 | ≥ 8500 |
| Osteoporosis | *≤ − 2.5* | *≤ 80 mg/cm$^3$* | *≤ − 2.5* | *≤ 80 mg/cm$^3$* |
| Low bone mass (osteopenia) | > − 2.5 to < − 1.0 | > 80 to < 120 | > − 2.5 to < − 1.0 | > 80 to < 120 |
| Normal bone mass | ≥ − 1.0 | ≥ 120 | ≥ − 1.0 | ≥ 120 |

The strength thresholds were developed and validated for measurement by VirtuOst. Strength is reported to ± 10 N, BMD T-score to ± 0.1 unit, and volumetric BMD to ± 1 mg/cm$^3$. For the overall fracture risk classification (see text), a patient is classified as being at: *high risk* if testing test positive for fragile bone strength or (BMD-defined) osteoporosis or both, at either the hip or spine; *not increased risk* if both bone strength and BMD are in the normal range at all measured sites; or *increased risk* otherwise

analyses are performed in an expert and consistent manner across different technicians, CT scanners, acquisition settings, patients, and over time.

## Clinical interpretation of BCT results

By comparing a patient's measurements of bone strength and BMD to respective interventional thresholds (see below), BCT provides classifications for *fragile bone strength*, *low bone strength*, and *normal bone strength* and for (BMD-defined) *osteoporosis*, *low bone mass* (aka, *osteopenia*), and *normal bone mass* (Table 1). Based on those classifications, an overall fracture risk classification is assigned following traditional DXA criteria [46], expanded to consider also the bone strength measurements:

- *High risk*, if the patient tests positive either for *fragile bone strength* or (BMD-defined) *osteoporosis*, or both, at either the hip or spine
- *Increased risk*, if the patient is not at high risk and instead tests positive for either *low bone strength* or *low bone mass*, or both, at either the hip or spine
- *Not increased risk*, if both bone strength and BMD are in the *normal* range at all measured sites

As cleared by the FDA, these classifications can be used by a physician to diagnose osteoporosis and assess fracture risk. For example, the *International Society for Clinical Densitometry's* practice guidelines for initiating therapeutic treatment [23] recommend that patients be considered for therapeutic treatment if classified as having fragile bone strength. As discussed in more detail below, results from BCT do not need to be confirmed by DXA, although some physicians may decide to subsequently order a baseline DXA for use in monitoring a treatment response.

Using these criteria, a slightly greater number of patients will test positive, and presumably more will be treated, than if one were to define high-risk patients on the basis of fragile bone strength alone or BMD alone. When using *only hip BCT*, the prevalence of high-risk patients by BCT appears to be similar to the prevalence of BMD-defined osteoporosis based on traditional DXA criteria (BMD T-score ≤ − 2.5 *at the hip or spine*). For example, in the large real-world FOCUS study of hip fractures [20] (the "**F**racture, **O**steoporosis, and **C**T **U**tilization **S**tudy", see below for details), prevalence of high-risk women by hip BCT using the above criteria was 30%, which was similar to the prevalence of 28% by traditional (hip/spine) DXA criteria (for patients not treated previously with osteoporosis medications). Based on these data, the number of positive testing patients should be similar if hip BCT is used instead of traditional DXA. As discussed below in more detail, the FOCUS study also demonstrated that these high-risk women (by hip BCT) were at significantly higher risk of hip fracture than those identified by traditional (hip/spine) DXA.

If BCT is performed at *both* the hip and spine, more high-risk patients by the above criteria will be expected than if using only hip BCT. One small study on CT colonography patients (age 43–92 years) reported the prevalence of high-risk patients when using both hip and spine BCT as compared to prevalence of BMD-defined osteoporosis by traditional DXA (hip/spine) [47]. That study reported prevalence of 33% by hip/spine BCT versus 27% for hip/spine DXA for 106 women, suggesting that use of hip/spine BCT will increase the number of positive-testing patients versus traditional (hip/spine) DXA. Additional and larger studies are required to characterize the risk profile of those patients. However,

since about 60–70% of abdominal scans (at least in the USA) are performed with intravenous contrast, which precludes accurate analysis of the spine [44], the issue of performing both hip and spine BCT will not arise for most patients in typical clinical practice. That said, sagittal reconstructions from abdominal scans with intravenous contrast can be used to identify prevalent vertebral fractures in the lumbar region [48, 49]. The clinical utility of combining such measurements with hip or spine BCT or both remains to be investigated.

Occasionally, some patients classified by BCT as high risk will have fragile bone strength without having BMD-defined osteoporosis. Mechanistically, such patients might have small bones, a relatively porous trabecular but normal cortical compartment [50], unusual 3D geometry, unusual spatial distribution of bone density, focally thin cortices [51], or weak subregions. These features are typically not reflected by the DXA BMD T-score but can be captured in the less averaged and more mechanistic finite element analysis, explaining why BCT-based bone strength has been found to predict incident fractures independently of BMD at both the spine [33, 35] and hip [20, 35]. Supported by the clinical evidence discussed below, these patients with fragile bone strength are classified as being at high risk of fracture regardless of the BMD measurement. Less commonly, some patients may be classified to have BMD-defined osteoporosis without fragile bone strength. As per traditional but well-established clinical guidelines, because these patients have BMD-defined osteoporosis, they are also classified as being at high risk of fracture regardless of the bone strength measurement. Patients testing positive for *both* fragile bone strength *and* BMD-defined osteoporosis are at additionally increased risk [20].

## Typical uses of BCT

To date, measurement of BMD by DXA has been the standard of care for osteoporosis testing for all patients. Going forward, based on the evidence reviewed below, BCT could now be considered as an accurate alternative to DXA for the following two situations:

### For opportunistic use

BCT would be appropriate for a patient who has had a recent CT or is about to undergo a CT for any medical indication and who satisfies both of the following criteria:

- would benefit medically from an accurate osteoporosis test and who meets clinical guidelines for osteoporosis testing
- has had a hip- or spine-containing CT scan for any indication (for hip: any abdominal or pelvic CT, including whole-body CT; for spine: any CT without intra-venous

contrast containing one or more lumbar or lower thoracic vertebrae)

For this situation, the patient does not need to undergo any extra imaging procedure for diagnostic purposes and there is no extra radiation exposure because the patient's CT scan has already been acquired or ordered for other reasons. The FOCUS study [20] from Kaiser Permanente Southern California established the robustness of this opportunistic approach for BCT, sampling all available hip-containing CT scans from a cohort of 111,694 patients, acquired on 80 different CT scanners, across 14 hospitals, over a 9-year time frame. Approximately two-thirds of the CT scans were acquired using intravenous contrast, and more than one-third of CT scans had a slice thickness between 3 and 5 mm. In total, 86% of all available hip-containing CT scans were analyzable by BCT, the most common reason for rejection being insufficient distal bone coverage by the scan (more common in abdominal than in pelvic or pelvic-abdominal scans).

### For non-opportunistic use

BCT with a dedicated CT scan would be appropriate for patients for whom a DXA test is either not easily available, inadequate, or inconclusive. Examples include institutions or regions with CT but without DXA facilities; if appreciable bone deformities or degenerative changes exist that would compromise the accuracy of a DXA measurement; if the patient is highly obese, which can be particularly problematic for DXA [52, 53]; or if more detail is required for assessment purposes than is provided by DXA. An example of the latter would be a patient presenting with a wrist fracture who has low bone mass by DXA and for whom additional information on the hip or spine is required to make a diagnosis.

For opportunistic BCT, the patients' primary care physician or health care provider would typically order the BCT test at some time after the original CT was taken; the physician who ordered the original CT would typically not be involved. In addition, depending on the particular hospital system and regional practice guidelines, there are applications for which the same physician who orders the original CT scan may simultaneously order BCT. Examples include (Table 2) patients undergoing the following: CT enterography for assessment of inflammatory bowel disease (IBD) [45], CT colonography for colorectal cancer screening [47], spine CT before spinal fusion surgery [55], or PET/CT for staging prostate cancer [56]. In these cases, the CT is typically ordered to help manage the underlying medical condition and the guidelines for managing that condition also suggest bone density testing due to the medical condition itself or associated medications, either of which is associated with deteriorated bone strength. For the busy specialist physician and imaging-burdened patient typical of these medical conditions, ordering and undergoing

**Table 2** Use of BCT in clinical studies by related clinical application

| Study | Sample size by sex | Clinical application | Type of CT scan |
|---|---|---|---|
| Weber 2014 [45] | 71 F; 65 M | Inflammatory bowel disease | CT enterography |
| Maldonado 2019 [54] | 135 F; 122 M | Inflammatory bowel disease | CT enterography |
| Fidler 2016 [47] | 136 F | Colorectal cancer screening | CT colonography |
| Burch 2016 [55] | 98 F | Spine fusion | Pre-operative spine CT |
| Schwaiger 2017 [56] | 82 M | Prostate cancer | PET/CT |
| Adams 2018 [20] | 1,922 F; 861 M | Opportunistic screening | Any hip-containing CT |

*F* female, *M* male

opportunistic BCT may be less taxing and therefore more appealing to both patient and provider compared to arranging for and undergoing a separate DXA test. Early experience for patients with inflammatory bowel disease indicates this opportunistic approach can indeed lead to greater compliance with the clinical guidelines for bone testing [54].

Currently, opportunistic BCT is best suited for diagnostic purposes as opposed to monitoring a patient's response to treatment. For the latter, detecting a statistically significant treatment response—or a lack of response—over a 1–2-year period for an individual patient requires the use of the same acquisition settings and scanner manufacturer for the serial CT scans. Because this is typically difficult to achieve for opportunistic BCT, we recommend that for patients tested by BCT opportunistically, a baseline DXA could be ordered to monitor the treatment response. In this way, opportunistic BCT could be used for diagnostic purposes and DXA for monitoring a treatment response; if a patient tests negative by BCT, they can be tested diagnostically again at some later time using either BCT or DXA or any appropriate osteoporosis test.

## Clinical efficacy and validation

In this section, we summarize the evidence supporting the VirtuOst implementation of BCT. Key clinical points are first presented, each of which is then justified by the accompanying discussion. Additional support on each topic is presented in a series of detailed appendices (see Supplementary Materials).

### The BMD measurements from BCT can be used to identify osteoporosis and assess fracture risk using traditional clinical guidelines and FRAX or other risk calculators

One key issue when measuring BMD for clinical decision-making is to ensure proper calibration of the CT scan, particularly with opportunistic BCT for which an external calibration phantom is not used. The VirtuOst implementation of opportunistic BCT uses a patient-specific phantomless calibration, in which the patient's own internal tissues (e.g., blood,

visceral fat) and air—all assessed from the patient's CT scan—are used as calibrating references [32]. If there is excessive image noise or metal artifact throughout these reference tissues, the tissues cannot be used for internal calibration. As discussed below, four blinded, prospective clinical studies have reported the validity of this general approach for measuring BMD [20, 32, 45, 47].

For the hip, BCT provides DXA-equivalent BMD T-scores that can be used with the NHANES database of reference values. The approach BCT uses for BMD T-scores is similar to what is used by contemporary Lunar DXA machines, in which Lunar-measured BMD values are mapped into Hologic-equivalent values using empirical relations [57]. That mapping then enables the Hologic-measured young-reference values from the NHANES cohort to be used with a Lunar DXA machine when calculating NHANES-compatible T-scores—although the Lunar machine was not used on the NHANES cohort. In the same way, when calculating T-scores, the BMD values from VirtuOst are mapped to Hologic-equivalent values to enable use of the NHANES reference values.

To validate that approach for opportunistic BCT, hip areal BMD T-scores derived separately from opportunistic BCT and DXA were compared in two different studies in which the BCT analyses were performed blinded to the DXA data. In each study, the patients had both DXA and a pelvic-abdominal CT exam as part of their routine clinical care—either a CT enterography with intra-venous contrast ($n = 65$ men, 71 women; age 18–85 years) [45] or a low-energy CT colonography without intra-venous contrast ($n = 136$ women; age 43–92 years) [47]. In both studies, there was a high correlation between the femoral neck BMD T-score as derived from DXA versus BCT ($R^2 = 0.84$) and good agreement in an absolute sense as demonstrated by Bland-Altman analyses. In the CT colonography study, for example, the BMD T-scores for BCT agreed with DXA (both for the left hip) to the same extent as did the BMD T-scores by DXA between the left vs. right hips (Fig. 2). Further, in the colonography study, all eight patients with BMD-defined osteoporosis by DXA had BMD-defined osteoporosis by BCT (sensitivity 100%, specificity 98%) [47], as did six of the seven patients in the enterography study (sensitivity 86%, specificity 97%) [45]. These collective
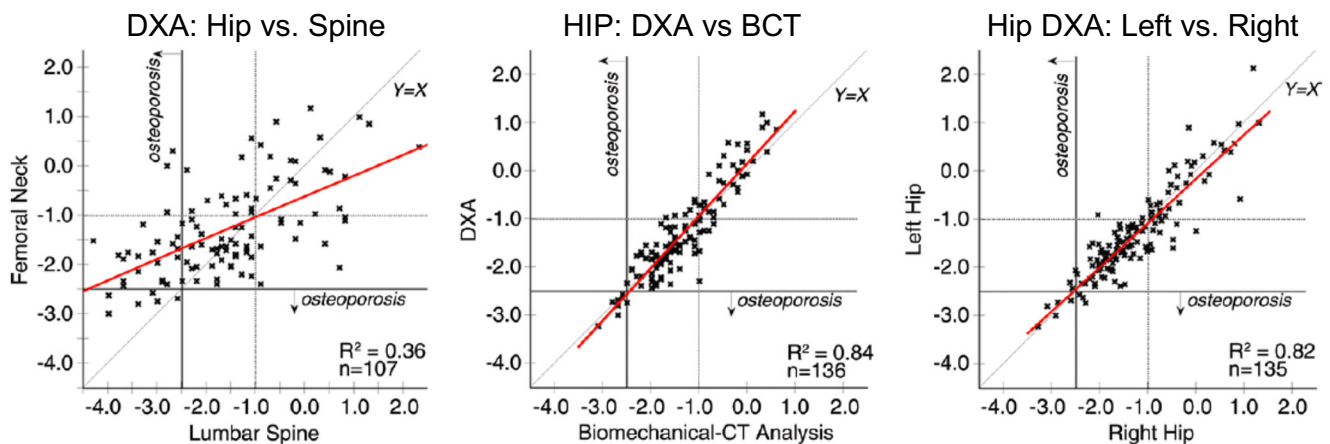
Fig. 2 Comparison of BCT vs. DXA for measuring the femoral neck BMD T-score, with various DXA comparisons for reference. The correlation between BCT and DXA (center) was similar to that between left and right hip for DXA (right) and over twofold higher than between hip and spine for DXA (left). From Fidler [47]

findings demonstrate that the femoral neck BMD T-scores for BCT and DXA in real-world practice are statistically equivalent. Other groups have also validated DXA-like hip areal BMD measurements obtained from calibrated CT scans [58–60]. At present, BCT only assesses one hip; it remains to be seen if there is any clinical utility in assessing both hips.

As would be expected from this statistical equivalence, the hip BMD T-scores from BCT and DXA have also been shown to be equally effective for assessing fracture risk and predicting hip fracture. Sampling from over 110,000 patients who had both DXA and abdominal CT as part of their medical care, the FOCUS study [20] determined the association between the hip BMD T-score (lower value from the femoral neck and total hip regions), obtained from both BCT and DXA, against the occurrence of new hip fractures subsequent to the CT and DXA scans. The BCT measurements were made blinded to the DXA and fracture-outcome data. Results indicated that for all the women tested with BCT in that study—1019 with hip fracture and 903 without—the mean values of the hip T-score from BCT and DXA differed by at most 0.1 T-score unit. Further, the age-adjusted hazards ratio per standard deviation deficit for hip fracture was the same for both modalities (HR/SD = 2.1), as was the AUC (0.72). Consistent with those metrics, at the traditional BMD T-score interventional threshold of − 2.5, specificity for predicting new hip fracture by the hip BMD T-score was the same for BCT and DXA (0.77), and sensitivity was similar (0.56 BCT, 95% CI 0.51–0.60; 0.52 DXA, 0.47–0.56). For the men, the sensitivity (0.45 BCT vs. 0.43 DXA) and specificity (0.82 BCT vs. 0.83 DXA) were similar between BCT and DXA, as was AUC (0.71 BCT vs. 0.73 DXA). Finally, hazard ratio for hip fracture for those testing positive by the BMD T-score criterion was statistically similar for BCT and DXA, both for women (3.7 BCT vs. 2.9 DXA) and men (4.0 BCT vs. 3.3 DXA).

Because both DXA and BCT use the NHANES III reference populations for calculating the hip BMD T-scores, one

implication of these collective findings is that the clinical guidelines for interpreting hip T-scores by DXA can also be used for BCT; in addition, the BCT-derived femoral neck BMD T-score can be used in lieu of the DXA BMD T-score in FRAX [28] or other risk calculators.

As noted earlier, the spine BMD measurement for BCT is not DXA-equivalent but instead is a volumetric BMD measurement for the trabecular bone within one vertebral body, thus avoiding many of the degenerative changes in the spine that can confound a DXA-type BMD measurement. Obtained for opportunistic BCT without an external calibration phantom, these BMD measurements have been validated in a study that directly compared them against paired measurements obtained by traditional quantitative CT [32]. That study utilized measurements derived from multiple clinical research studies that used traditional quantitative CT (and an external calibration phantom). A paired comparison was made of data from opportunistic BCT versus quantitative CT for 25 women and 15 men (age 41–86 years) scanned using 24 different CT scanners (from four different CT manufacturers). Results indicated a negligible difference (1 mg/cm$^3$, not statistically significant) between the two paired measurements and a high correlation ($R^2 = 0.98$, slope not different than unity); Bland-Altman analysis also revealed no bias. These results establish that the spinal volumetric trabecular BMD measurement from opportunistic BCT with internal tissue-based phantomless calibration is equivalent to that from traditional quantitative CT with an external calibration phantom; a similar level of agreement between the phantom and phantomless measurements was found for all the BCT measurements [32].

The clinical use of vertebral trabecular BMD for osteoporosis and vertebral fracture risk assessment, reported first in the 1980s [61], is supported with BCT by its consistently high association with vertebral fracture, using DXA as a reference. For example, in all five of the vertebral fracture-outcome studies discussed below for BCT, age-adjusted odds ratios per

standard deviation for vertebral fracture were consistently higher for vertebral trabecular BMD than for spinal DXA BMD, both for prevalent fracture—1.9 vs. 0.7 [62]; 1.9 vs. 1.3 [63]; 3.4 vs. 1.9 [64]—and for incident fracture—5.7 vs. 3.2 [33]; 2.4 vs. 1.8 [34]. Limited comparative data exist on sensitivity and specificity for predicting new vertebral fractures by vertebral trabecular BMD versus spinal DXA BMD. As discussed below, the available comparative data [33–35] suggest that the sensitivity for vertebral trabecular BMD at the ACR-recommended (American College of Radiology) cut-point of 80 mg/cm$^3$ is higher than for spinal DXA BMD T-score at the traditional $T \leq -2.5$ cut-point, although specificity can be lower for BCT. Other groups previously validated vertebral trabecular BMD from quantitative CT for vertebral fracture risk assessment [61]. Taken together, these studies demonstrate that the vertebral trabecular BMD measurement from opportunistic BCT is equivalent to that from traditional quantitative CT, which in turn is at least as good as spinal DXA-BMD for assessing vertebral fracture risk.

BCT does not currently use any volumetric measurement of BMD at the hip, in part because some fracture-outcome studies have shown no advantage of doing so over traditional hip BMD by DXA [65] and in part because interventional thresholds for volumetric measurements of BMD at the hip have not been established or validated.

## Bone strength by BCT accurately measures bone strength in human cadavers and has accurately quantified treatment effects on bone strength in monkeys

BCT is the only clinical test that non-invasively measures bone strength—in units of force—for fracture risk assessment. The bone strength measurement in BCT simulates cadaver-lab testing conditions, in which an excised bone is loaded to failure in a controlled orientation and configuration in order to measure the breaking force (strength) of the bone. Extensive literature reviews of BCT in general for bone strength assessment are available elsewhere [24, 26]. For the VirtuOst implementation of BCT or its earlier versions, four studies have reported on the accuracy of BCT-derived measurements of bone strength, three studies addressing vertebral strength (for a compressive overload) [33, 66, 67], and one addressing femoral strength (for a sideways fall) [14]. All four studies used laboratory-based biomechanical testing as the gold standard. As reviewed in detail in Appendix A, these human cadaver studies consistently demonstrated that BCT accurately measured bone strength at the hip (sideways fall) and spine (compressive overload) and that the directly measured strength from biomechanical testing was more highly correlated with BCT-measured bone strength than BMD (either by DXA or quantitative CT) [14, 33, 66]. Consistent with these

results for VirtuOst, other groups using different implementations of BCT have also reported correlations with directly measured cadaver bone strength that were higher for BCT-measured bone strength than BMD, both for the spine [68–70] and hip [70–74].

All of these validation studies were done in human cadaveric specimens with limited information about prior medication use. To address the question of whether BCT measurements are accurate in the setting of osteoporosis therapies, the VirtuOst implementation of BCT was applied to non-human primate bones with and without drug treatment (with denosumab). In that study, BCT accurately measured vertebral strength both with and without drug treatment and with and without ovariectomy [67]. Vertebral strength by BCT was highly correlated with gold-standard biomechanical testing ($R^2 = 0.97$, $n = 52$ monkeys; Fig. 3), the relation being independent of treatment ($p = 0.12$). Further, the magnitude of the treatment effect of + 51% (95% CI 20–88%) observed with BCT was just slightly numerically lower than the treatment effect of + 57% (95% CI 26–95%) by biomechanical testing (treatment effect estimated by comparing treated vs. untreated groups). By contrast, the magnitude of the treatment effect by bone mineral content and volumetric BMD of the vertebral body was over twofold lower at 27% (95% CI 8–50%) and 28% (95% CI 14–45%), respectively.

In addition to demonstrating that BCT accurately measured the magnitude of the treatment-induced effects in bone strength in this monkey model, these findings also demonstrate the complexity of interpreting the magnitude of the
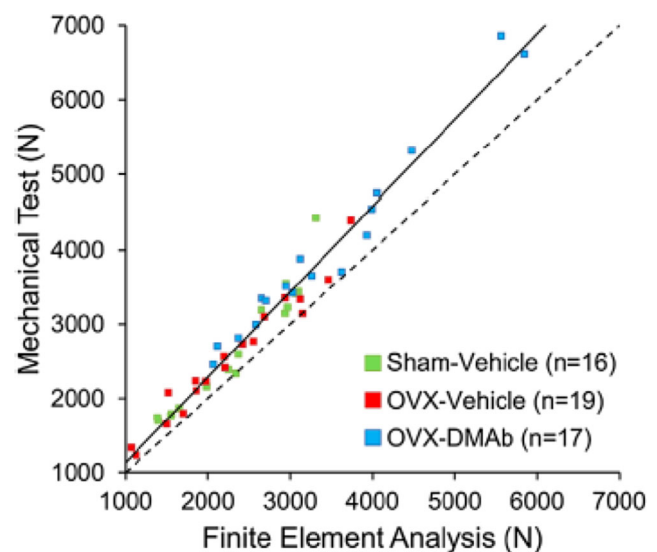


**Fig. 3** Prospective validation of BCT (finite element analysis) for measuring T12 vertebral strength in aged monkeys with and without drug treatment, using direct mechanical testing as the gold standard. BCT was performed blinded to the mechanical test results. Data are color-coded by treatment group, for 52 monkeys, some treated with denosumab or vehicle after ovariectomy, vs. sham ($R^2 = 0.97$, $p < 0.0001$; all data pooled); for reference, the dashed line shows the line of unity. From Lee [67]

treatment-induced changes in BMD or bone mass with respect to actual changes in bone strength. For the monkey study, the density/mass parameter that had a magnitude of effect most similar to the observed change in bone strength was the (volumetric) bone mineral content of the cortical shell and its thickness, as measured by micro-CT analysis [67]. However, biomechanically, BCT also showed that the strength associated with the cortical and trabecular compartments changed approximately equally [67]. It is not possible to directly validate these findings in humans. Short of that direct validation, the data so far suggest that the main determinants of treatment effects on strength—at least in the monkey model and for treatment with denosumab—are those that were adequately captured by BCT, namely, the bone geometry and mass and the spatial distribution of bone mass. Those findings in turn imply that any potential molecular or other lower-scale effects of the treatment in that monkey study did not play any appreciative role in the strength response. A similar trend was observed in a finite element analysis study of the distal radius in monkeys treated with odanacatib [75]. If molecular or small-scale treatment effects are also unimportant after treatment with contemporary osteoporosis agents in older humans, these collective human cadaver and monkey validation studies suggest that BCT may capture the correct magnitude of treatment effects on bone strength in humans.

## Bone strength by BCT is associated with risk of hip, spine, and major osteoporotic fractures at least as strongly as is BMD by either DXA or quantitative CT, for both sexes

Nine clinical fracture-outcome studies have been performed to date for BCT using VirtuOst. Involving both women and men, three studies addressed prevalent spine fracture, two addressed incident spine fracture, two addressed incident hip fracture, one addressed both incident spine and hip fractures, and one addressed any prevalent major osteoporotic fracture (clinical spine, hip, proximal humerus, or wrist). Together, these studies involved BCT measurements taken in over 5500 subjects, sampled using case-control or case-cohort designs from much larger study populations. In all instances, the BCT analyses were performed blinded to the fracture outcomes. Details of the studies and cohorts are provided in Appendix B. Overall, the combined cohorts represented both sexes, populations in both the USA and Europe, with one large US cohort [20] including both sexes and being racially diverse.

Across all nine studies, low values of bone strength were consistently associated with an increased risk of fracture, including hip, spine, or any osteoporotic fracture, for both sexes (Table 3; Fig. 4). As a reference for interpretation, results for bone strength by BCT in these studies can be compared with those for BMD, usually by DXA (or CT-based DXA-

equivalent measurements). As noted elsewhere [77], an increase in the hazards or odds ratio (divided by the population standard deviation of the measurement) typically improves sensitivity for fracture prediction without markedly affecting specificity, although, as noted below, sensitivity and specificity data comparing both BCT and BMD are sparse. We did not perform a statistical meta-analysis of these data, which was beyond the scope of this review. However, collectively in 13 of the 14 comparisons of fracture outcomes made in these nine studies, the published data for each individual study indicate that bone strength had either a statistically stronger or numerically higher association with fracture than did BMD, as quantified by the age-adjusted hazard or odds ratio (divided by the standard deviation); the other single comparison showed an almost identical association. For example, across studies, the hazard or odds ratio per SD for spine fracture ranged from 1.7 to 7.2 for vertebral strength and from 0.7 to 3.2 for BMD and for hip fracture ranged from 3.0 to 8.0 for femoral strength and from 2.3 to 4.6 for BMD (Table 3; Fig. 4). Taken together, these data indicate that the association of bone strength with risk of hip or spine fractures equals or exceeds that of BMD by either DXA or quantitative CT.

In eight of these nine studies, BCT was performed at the same general anatomic site (hip or spine) as the site of the outcome fracture (hip or spine fracture, respectively). In the other study, hip BCT was used to assess risk of any major osteoporotic fracture [76]. That study demonstrated that femoral strength by hip BCT was associated with any (prevalent) major osteoporotic fracture, performing as well as DXA-equivalent hip BMD (by BCT). In particular, the age-adjusted odds ratio of a major osteoporotic fracture was at least as high for femoral strength as for the hip BMD (by BCT), both for both women [odds ratio (95% CI)—bone strength 1.8 (1.1–2.9) vs. hip BMD 1.5 (1.1–1.9)] and men [bone strength 3.2 (1.7–6.2) vs. hip BMD 2.0 (1.3–3.0)].

These collective data indicate that the odds or hazard ratio (per standard deviation) trended consistently higher for bone strength than BMD. Presumably, this trend reflects that a greater portion of the variation across the population in true bone strength—which directly influences fracture risk—is captured by the BCT measurement of bone strength than by BMD. Across large groups of individuals, the relative variation (ratio of standard deviation to mean value) for both femoral [20] and vertebral [63] strength by BCT is up to twofold greater than the relative variation for BMD, particularly for areal BMD. This greater relative variation for strength partly reflects the non-linear relation between bone strength and BMD at the tissue level; it also partly reflects other factors that affect whole-bone strength independently of BMD and that also vary across the population, for example, cortical thickness, overall bone shape and geometry, and spatial distribution of BMD within the bone including the relative amount of cortical and trabecular bone—all of which are captured to

**Table 3** Summary of studies comparing fracture risk assessment between BMD and bone strength

| Study by fracture site | Demographic N fracture/no-fracture, sex (N = 5514 total) | | Age-adjusted HR (per SD) or OR (95% CI) | |
|---|---|---|---|---|
| | | | BMD | Bone strength |
| Vertebral fracture (BCT/DXA measurements made at the spine): | | | | |
| Melton 2007 [62][†] | 40/40 | F | 0.7 (0.4–1.2) | 2.2 (1.2–4.3) |
| Melton 2010 [63][†] | 193/90 | F | 1.3 (1.1–1.6) | 1.7 (1.4–2.1) |
| Wang 2012 [33] | 63/243 | M | 3.2 (2.0–5.2) | 7.2 (3.6–14.1) |
| Anderson 2014 [64][†] | 40/80 | F/M | 1.9 (1.1–3.3) | 3.0 (1.4–6.9) |
| Kopperdahl 2014 [35] | 117/380 | F | 2.3 (1.7–3.2)[††] | 2.8 (1.8–4.3) |
| | 50/296 | M | 1.7 (1.2–2.5)[††] | 2.2 (1.5–3.2) |
| Allaire 2018 [34] | 26/62 | F/M | 1.8 (1.0–3.3) | 3.8 (1.5–9.2) |
| Hip fracture (BCT/DXA measurements made at the hip): | | | | |
| Orwoll 2009 [16] | 40/210 | M | 4.6 (2.6–8.3) | 8.0 (2.6–24.3) |
| Kopperdahl 2014 [35] | 108/500 | F | 2.7 (1.9–3.9) | 4.2 (2.6–6.9) |
| | 63/377 | M | 3.7 (2.5–5.6) | 3.5 (2.3–5.3) |
| Adams 2018 [20][§] | 595/769 | F | 2.3 (1.9–2.9) | 3.6 (2.8–4.6) |
| | 308/404 | M | 2.4 (1.9–3.1) | 3.0 (2.2–4.3) |
| Any major osteoporotic fracture (BCT/DXA measurements made at the hip): | | | | |
| Amin 2011 [76][†] | 55/175 | F | 1.5 (1.1–1.9) | 1.8 (1.1–2.9) |
| | 28/162 | M | 2.0 (1.3–3.0) | 3.2 (1.7–6.2) |

Incident (new) fractures unless noted otherwise; BMD is measured by DXA (or is DXA-equivalent) unless noted otherwise; bone strength is measured by VirtuOst or earlier implementations

*F* female, *M* male

[†] Prevalent (existing) fracture (occurred before the imaging)

[††] Vertebral trabecular volumetric BMD by quantitative CT (spine BMD by DXA not measured in those studies)

[§] Results are reported here for patients not recently treated with any osteoporosis medication

some degree and mechanistically within the BCT model but are missed by the DXA BMD measurement. For the same reasons, a simultaneous variation of these other factors over time also explains why typical age-related declines are thought to be greater (percent-wise) at the whole-bone level for bone strength than for BMD [78]. As discussed next, integration of these same factors in the BCT model likely explains why bone strength by BCT has been shown to predict fracture independently of BMD.

## Bone strength by BCT is associated with fracture risk independently of BMD by DXA

The large FOCUS study provided definitive evidence on this issue regarding hip fractures, for both sexes. That study drew from an underlying population of 111,694 patients age 65 or older with an abdominal or pelvic CT and a DXA exam within 3 years of the CT and no prior hip fracture. Cases were defined as those who subsequently suffered a hip fracture (1340 women, 619 men) and controls (no fracture) were randomly selected subgroup from the overall cohort (1350 women, 629 men), after removing any participants with hip fracture. As noted

above, this study confirmed the equivalence of the hip T-score (using the lower T-score from the femoral neck and total hip regions) as measured by BCT versus DXA. This equivalence justified using logistic regression to directly compare paired measurements of the hip BMD T-score (from BCT) and bone strength, both measurements taken from the very same CT scan and thus minimizing any random errors associated with variations in imaging parameters or any time difference between the CT and DXA scans.

FOCUS established that incident hip fracture was associated with femoral strength independently of hip BMD. In particular, after adjusting for age, BMI, race/ethnicity, and the hip BMD T-score (lower T-score from femoral neck or total hip regions, from BCT), the age-adjusted hazard ratio per standard deviation for femoral strength was statistically significant for both women (HR = 1.9, 95% CI 1.2–3.1) and men (HR = 3.4, 95% CI 1.6–7.2). Consistent with these findings, the earlier but smaller Age, Gene/Environment Susceptibility (AGES) study of hip fracture (108 fracture cases for women, 63 for men) also used logistic regression to show that femoral strength was associated with incident hip fracture independently of femoral neck areal BMD (by BCT) for women (p = 0.01) and independently of total hip areal BMD for both
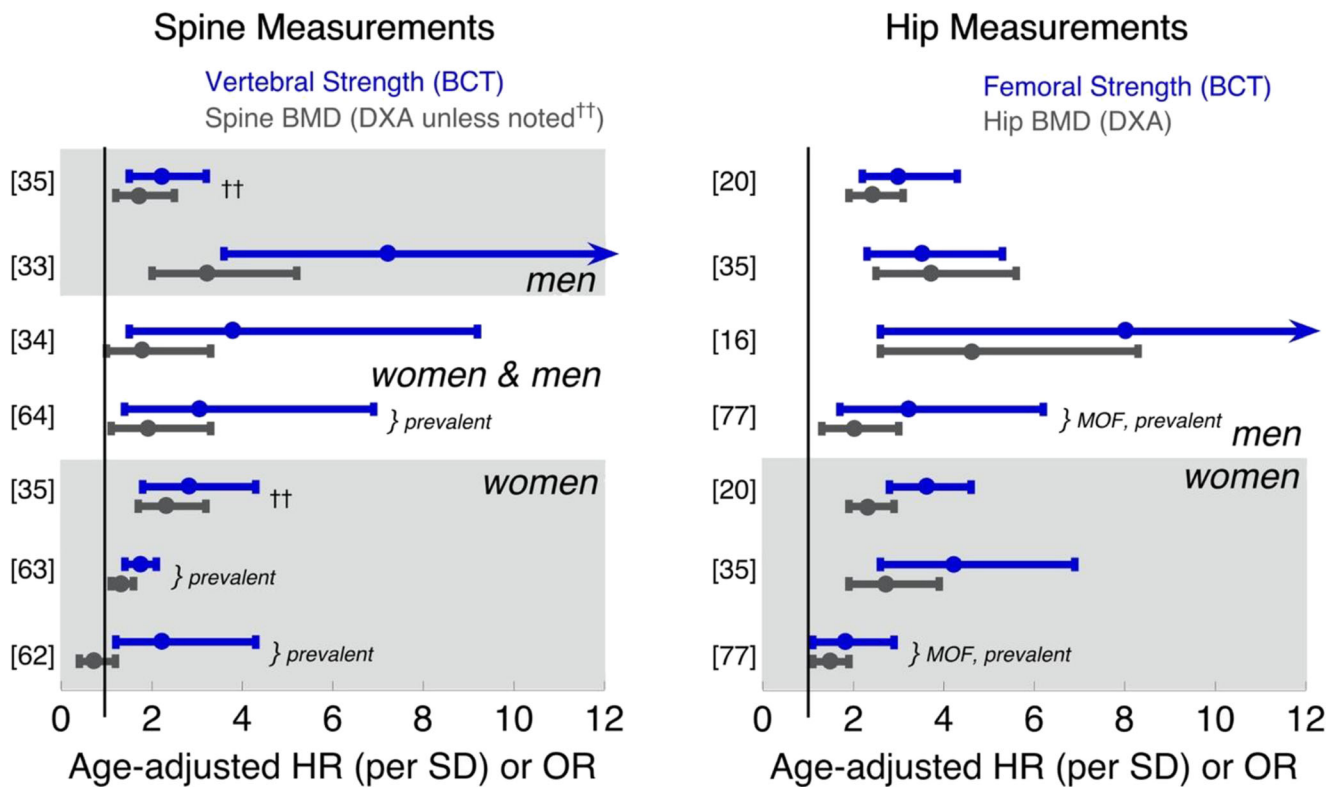
**Fig. 4** Graphic depiction of the age-adjusted hazard ratio (per standard deviation change) or odds ratio values, taken from Table 3, for hip and spine measurements, grouped by sex, for predicting fracture [study citations]. For the hip measurement, the outcome was incident (new) hip fracture, unless noted otherwise as prevalent (existing) MOF (major osteoporotic fracture). For the spine measurement, the outcome was incident vertebral fracture, unless noted otherwise as prevalent. BMD either by DXA or DXA-equivalent, unless noted (double dagger denotes trabecular BMD from quantitative CT). Error bars show the reported 95% confidence intervals (see Table 3 and Appendix B for details)

sexes (women $p = 0.0006$, men $p = 0.0001$) [35]. Previously, the slightly smaller Osteoporotic Fractures in Men (MrOS) study of hip fracture in men ($n = 40$ hip fractures) [16] did not show a significant independent association of femoral strength over hip BMD (HR = 2.7, 95% CI 0.5–14.6). However, the effect size in that study (HR = 2.7) is similar to that observed for men in the larger FOCUS study (HR = 3.4), suggesting the lack of statistical significance in the MrOS study was due to low statistical power.

Evidence also shows that vertebral strength is associated with risk of incident spine fracture independently of BMD by either spine or hip DXA. For the MrOS study of elderly men with clinically apparent new vertebral fractures ($n = 63$ fractures) [33], the age-adjusted hazard ratio per standard deviation deficit for vertebral compressive strength (7.2; 95% CI 3.6–14.1) was over twofold higher ($p < 0.005$) than for DXA lumbar spine BMD (3.2, 2.0–5.2) and was fourfold higher than for femoral neck BMD (1.8, 1.2–2.9). Further, vertebral strength was associated with vertebral fracture independently of DXA spinal BMD ($p < 0.001$) and its hazard ratio was numerically higher than for volumetric BMD of the entire vertebral body by quantitative CT ("integral BMD," 5.7, 3.1–10.3). In a more recent but smaller study of incident vertebral fracture in both sexes (13 men and 13 women with

fracture) [34], vertebral strength was associated with fracture independently of a research-only measurement of DXA-equivalent spinal BMD (age-adjusted odds ratio 5.1, 95% CI 1.5–17) and also had better prediction (AUC = 0.80 vs. 0.72, $p = 0.05$). The AGES study of incident vertebral fracture in Iceland ($n = 117$ women and $n = 50$ men with fracture) did not include DXA but did include volumetric trabecular BMD by quantitative CT. Consistent with the findings from MrOS for clinical vertebral fractures in men, vertebral strength was associated with fracture independently of volumetric BMD for the men ($p < 0.01$), and the combination of vertebral strength and vertebral trabecular volumetric BMD significantly improved ($p < 0.006$) the net fracture classification index for moderate/severe grade incident vertebral fractures in men. In that study, no significant improvement for vertebral strength over volumetric BMD was observed for predicting less severe vertebral fractures or for women, but the age-adjusted odds ratios per standard deviation all trended higher for vertebral strength than for volumetric BMD for both sexes and all fracture grades (e.g., 4.3 vs. 3.1 for women, 2.4 vs. 1.4 for men, for more severe fractures). Taken together, this body of evidence suggests that vertebral strength is associated with risk of *incident* spine fracture independently of BMD by DXA, with a slightly stronger association than seen for volumetric BMD

by quantitative CT. For *prevalent* fracture, the association with fracture is similar between vertebral strength and volumetric BMD, both of which consistently have stronger associations than DXA BMD [62–64].

## Older patients classified by BCT as having either BMD-defined osteoporosis or fragile bone strength are at high risk of fracture

Although fracture risk in older individuals depends on many factors, almost all clinical decision-making for considering therapeutic treatment ultimately classifies the individual patient into a "high-risk" category using some type of interventional threshold. For example, the DXA BMD T-score threshold of −2.5 at the hip or spine is widely used to identify candidates for treatment. Likewise, in the USA, the 3% threshold for 10-year absolute risk of hip fracture from the FRAX calculator is also used to identify high-risk patients suitable for treatment [46, 79]. Research on optimal ways to identify high-risk patients using some type of threshold or risk-based approach, and how to best incorporate strength measurements, is ongoing in the field.

For BCT, the FDA-approved interventional thresholds for fragile bone strength using VirtuOst, first reported in 2014 [35], were developed with clinical decision-making in mind and have been validated in a number of studies (for details on the development, see Appendix C). To validate that patients classified by BCT as having fragile bone strength are indeed at clinically significant high risk of fracture—and therefore candidates for treatment—comparisons of sensitivity (and specificity) for predicting new fractures via fragile bone strength by BCT can be compared with sensitivity (and specificity) via BMD-defined osteoporosis by DXA, the clinical standard of care (or DXA-equivalent hip BMD from BCT); alternatively, the observed probability of fracture in these studies can also be compared at the interventional thresholds for fragile bone strength versus BMD-defined osteoporosis, again interpreting the latter as a reference standard. Two such studies addressed new hip fractures [20, 35], two addressed new spine fractures [34, 35], and two earlier studies also provide support when considered in retrospect [16, 33]; details of all studies are provided in Appendix C. Collectively, the available data demonstrate that older patients classified by BCT as having fragile bone strength, at the hip or spine, are at clinically significant high risk of fracture.

As regards to identifying patients at high risk of *hip* fracture, the FOCUS study assessed both risk of fracture at the interventional thresholds and sensitivity (and specificity) [20]. Risk of hip fracture was numerically higher for the women and men who tested positive with fragile bone strength by BCT than for those testing positive with BMD-defined osteoporosis either by hip DXA or by traditional DXA (hip/spine), the difference against traditional DXA reaching statistical significance for the women,

and AUC values for femoral strength were similar to those from the DXA BMD T-score (within 0.10 points). Considering sensitivity and specificity, for the women in the FOCUS study, sensitivity for predicting incident hip fracture at 5 years was significantly higher for fragile bone strength at the hip by BCT (0.63, 0.59–0.68) than for BMD-defined osteoporosis at the hip by DXA (0.52, 0.47–0.56), although specificity was significantly lower for fragile bone strength (0.69, 0.64–0.74) than for BMD-defined osteoporosis (0.77, 0.73–0.81). For men, sensitivity trended numerically higher for fragile bone strength at the hip (0.48, 0.42–0.55) than for BMD-defined osteoporosis at the hip by DXA (0.43, 0.37–0.50), with similar specificities (0.82 vs. 0.83, respectively). In typical clinical practice, one would use the lower BMD T-score from the hip or spine when using DXA to identify patients with BMD-defined osteoporosis [46]. Doing so in FOCUS yielded a sensitivity for BMD-defined osteoporosis for women that increased from 0.52 to 0.59—numerically lower than 0.63 for fragile bone strength—while specificity decreased from 0.77 to 0.67—also numerically lower than 0.69 for fragile bone strength. As described in more detail in Appendix C, in the AGES study of incident hip fractures in women and men in Iceland [35], the elevated probability of fracture at the interventional thresholds was statistically similar for fragile bone strength versus BMD-defined osteoporosis at the hip, and retrospective evidence of higher sensitivity at equivalent specificity for fragile bone strength at the hip over BMD-defined osteoporosis at the hip was reported in the MrOS study of hip fractures in elderly men [16].

The FOCUS study extended these results by showing that when *both* femoral strength and hip BMD from BCT are used to identify high-risk patients—as opposed to using just femoral strength or just BMD—BCT at the hip correctly identified more patients at high risk of hip fracture than did traditional (hip/spine) DXA, and those patients testing positive with hip BCT were at higher risk of hip fracture than those testing positive by traditional DXA [20]. Consistent with that finding, in the earlier AGES study, reclassification analysis indicated that prediction of hip fracture for women was improved by considering both femoral strength and hip BMD ($p = 0.002$) [35].

Considering measurements of both femoral strength and hip BMD can facilitate clinical interpretation since it enables physicians to easily identify patients with clinically significant low levels of bone strength in the absence of BMD-defined osteoporosis. In doing so, one approach is to identify high-risk patients as those who test positive by BCT for *either* fragile bone strength *or* BMD-defined osteoporosis. Using this either/or approach in the FOCUS study, sensitivity for predicting 5-year hip fracture increased (0.63 to 0.66 for women; 0.48 to 0.56 for men) and specificity decreased (0.69 to 0.66 for women; 0.82 to 0.76 for men) compared to using only bone strength, and AUC values did not change. For clinical reference purposes, when comparing this approach for hip

BCT against traditional DXA (lowest T-score at the hip or spine), sensitivity for BCT was 12% higher for women (0.66 vs. 0.59) and 17% higher for men (0.56 vs. 0.48) than for traditional DXA, with similar values of specificity between hip BCT used in this way and traditional DXA. Furthermore, the women who tested positive by BCT in this way were at over 50% higher elevated risk of hip fracture than were those women testing positive by traditional DXA (hazard ratio 3.4 vs. 2.2, $p < 0.05$); there was no significant effect for men (hazard ratio 4.0 BCT vs. 3.3 DXA). Studies have not yet been performed to assess whether hip fracture prediction is improved for BCT by combining BCT measurements at the spine and hip. Nevertheless, these findings demonstrate that older patients classified by hip BCT as having either fragile bone strength or BMD-defined osteoporosis (by BCT) are at clinically significant high risk of hip fracture.

As regards to identifying patients at high risk of *spine* fracture, fewer data exist. As described in more detail in Appendix C, in the AGES study of incident vertebral fractures in women and men in Iceland [35], the elevated probability of fracture at the interventional thresholds was statistically similar for fragile bone strength versus BMD-defined osteoporosis at the spine; sensitivity and specificity were not reported. In the small Framingham study of women and men (26 incident fracture cases) [34], fragile bone strength at the spine trended toward twofold higher sensitivity than a validated DXA-equivalent spine T-score (from quantitative CT) for identifying new vertebral fracture (0.46 vs. 0.23, $p = 0.09$), at similar values of specificity. Consistent with these findings, in the earlier and larger ($n = 63$ incident fracture cases) MrOS study of clinical spine fractures in elderly men [33], low values of vertebral strength produced higher sensitivity than did the established spinal DXA T-score thresholds, at the same specificity. The interventional thresholds for fragile bone strength had not been established at the time of that study. Even so, sensitivity trended higher for strength both at 95% specificity (37% BCT vs. 30% DXA) and at 90% specificity (52% BCT vs. 43% DXA), and AUC was higher (0.83 BCT vs. 0.76 DXA, $p < 0.02$).

## BCT measurements are precise and clinically reproducible

Opportunistic BCT uses the patient's internal tissues as "phantomless" calibrating references; it also utilizes CT scans from a variety of different CT scanners and acquisition settings. One potential concern with opportunistic BCT is the measurement error associated with the use of different CT scanners or scan acquisitions—how robust are one-time measurements used for diagnostic purposes? To date, two studies have been reported that are relevant to measurement precision

and clinical reproducibility for opportunistic BCT (performed at the O.N. Diagnostics centralized BCT facility).

For inter-operator precision, one study was performed on 25 women and 15 men (age range, 41–86 years) who underwent CT scanning on 24 different CT scanners (four different CT manufacturers) as part of different clinical drug trials (baseline scans only, no drug treatment); scans were re-analyzed by two different BCT technicians who were blinded to each other [32]. Results indicated that the reanalysis precision errors (CV%) for all measurements of bone strength and BMD by opportunistic BCT (i.e., using phantomless calibration), at the hip and spine, were 0.5% or less. Thus, for the same scan analyzed by different BCT technicians, inter-operator discordance was negligible.

To assess clinical reproducibility, one must account for the typical uncertainty associated with the variability of the source scan, both in CT scanner and the acquisition settings, and patient repositioning. Data from one study can be used to address this issue [56]. That study reported on men with prostate cancer ($n = 82$, $71.6 \pm 8.3$ years) who underwent both PET/CT and multi-detector CT at different time points as part of their medical care—two very different types of CT scan. Both the BMD and bone strength measurements at the hip and spine were directly compared between the paired PET/CT (one scanner) and multi-detector CT (12 different scanners) scans, taken in a clinical setting within 3 months of each other (full paired data were available for $n = 63$ patients). Results indicated that the mean paired differences ($p > 0.05$ unless noted) between the various BCT measurements for the two types of scan were all small: 1.1% for total hip areal BMD, 1.3% femoral strength, 2.6% for vertebral trabecular BMD, 1.7% for vertebral strength, and 2.5% ($p = 0.007$) for femoral neck areal BMD. Consistent with these small differences, between-scan agreement for fracture-risk classification was 97% (0.89 kappa for repeatability). Comparable differences have been noted between different types of DXA scanners for the BMD T-score [80]. For example, for the femoral neck BMD T-score, which can be directly compared between BCT and DXA, the correlation between measurements within each modality was at least as high for BCT ($R^2 = 0.94$; PET/CT vs. multi-detector CT) as for DXA ($R^2 = 0.87$; Hologic DXA vs. Lunar DXA). These findings suggest that the clinical reproducibility for one-time measurements for opportunistic BCT when used on different scanners/settings should be comparable to reproducibility for DXA when used on different scanners.

## Non-opportunistic BCT can be used to monitor for a treatment response

"Monitoring" in osteoporosis care typically refers to measuring or confirming a *treatment-induced response* over time,

**Table 4** BCT monitoring studies (13 total) for drug or other treatments. Non-opportunistic BCT was performed at the hip (H) or spine (S) or at both sites. $N = 1644$ subjects total, over all studies. Details of the studies in italicized entries are discussed in Appendix D

| Study | Age (years) mean ± SD | Number | Hip or spine assessed | Treatment |
|---|---|---|---|---|
| Keaveny 2007 [82] | 62.5 ± 1.6 | 25 | S | Alendronate |
| | 64.5 ± 1.4 | 28 | S | Teriparatide |
| Mawatari 2008 [83] | 62.7 ± 2.2 | 15 | S | Control (non-antiresorptive) |
| | 60.4 ± 2.9 | 14 | S | Alendronate |
| Keaveny 2008 [84] | 55–85 | 72 | H | PTH(1–84) |
| | 55–85 | 42 | H | Alendronate |
| | 55–85 | 37 | H | PTH(1–84) and alendronate |
| Lewiecki 2009 [85] | 63.5 ± 6.0 | 46 | H + S | Placebo |
| | 64.8 ± 7.2 | 47 | H + S | Ibandronate |
| Keaveny 2012 [86] | 64.2 ± 7.7 | 27 | H | Teriparatide |
| | 62.2 ± 8.4 | 21 | H | Alendronate |
| *Brixen 2013 [87]* | *64.0 ± 6.2* | *105* | *H + S* | *Placebo* |
| | *63.9 ± 7.3* | *109* | *H + S* | *Odanacatib* |
| Cosman 2013 [88] | 67.9 ± 1.3 | 43 | H + S | Raloxifene, add teriparatide |
| | 68.6 ± 1.3 | 44 | H + S | Raloxifene, switch to teriparatide |
| | 67.1 ± 1.3 | 46 | H + S | Alendronate, add teriparatide |
| | 69.3 ± 1.2 | 49 | H + S | Alendronate, switch to teriparatide |
| Keaveny 2014 [31] | 74.1 ± 6.0 | 48 | H + S | Placebo |
| | 73.3 ± 4.2 | 51 | H + S | Denosumab |
| Kleerekoper 2014 [89] | 62 ± 8 | 34 | H + S | Teriparatide |
| *Snyder 2017 [90][†]* | *72.4 ± 5.5* | *101* | *H + S* | *Placebo* |
| | *72.3 ± 6.3* | *110* | *H + S* | *Testosterone* |
| Keaveny 2017 [91] | 64.3 ± 4.7 | 24 | H + S | Romosozumab |
| | 65.8 ± 5.7 | 31 | H + S | Teriparatide |
| | 66.1 ± 5.8 | 27 | H + S | Placebo |
| *Langdahl 2017 [92]* | *71.8 ± 7.4* | *218* | *H* | *Romosozumab (prior bisphosphonate)* |
| | *71.2 ± 7.7* | *218* | *H* | *Teriparatide (prior bisphosphonate)* |
| Howe 2019 [93] | 56–89 | 12 | H | AGN1 LOEP |

VirtuOst or a related software implementation of BCT was used in all studies

[†] Women in all studies except Snyder [90], in which only men were assessed

which can be challenging clinically. One challenge is that annual percent changes in bone strength (and BMD) for the individual patient are often comparable to measurement precision errors [81], and changes may be even smaller if the patient is on an antiresorptive treatment. Another challenge is that, for maximum measurement precision, the imaging equipment and protocol should remain unchanged across all serial measurements. Since those latter conditions are difficult to achieve for opportunistic BCT, and since *opportunistic* BCT has not yet been characterized for monitoring treatment responses in individual patients, monitoring of a treatment response for an individual patient is not currently recommended with opportunistic BCT. However, as discussed next, the evidence so far suggests that when BCT is used *non-opportunistically*, that is, with a dedicated CT scan and a traditional calibration phantom, and when the imaging

equipment and protocol remains unchanged across all serial measurements, then BCT is an excellent modality for monitoring treatment responses in individual patients and may detect changes missed by DXA.

The ability of non-opportunistic BCT to detect treatment responses is evident from multiple clinical research studies that assessed various types of treatment in which both BCT and DXA were used for monitoring. Using VirtuOst or its earlier implementations, 13 such clinical studies have been reported, involving over 1600 subjects (Table 4). Collectively, these studies demonstrate that loss of bone at the hip or spine in placebo groups can be detected earlier and is statistically more significant by BCT than by DXA (Fig. 5; see further details in Appendix D). Furthermore, some BCT studies have shown statistically significant changes in a strength-to-density ratio after drug treatment [82, 88, 89].
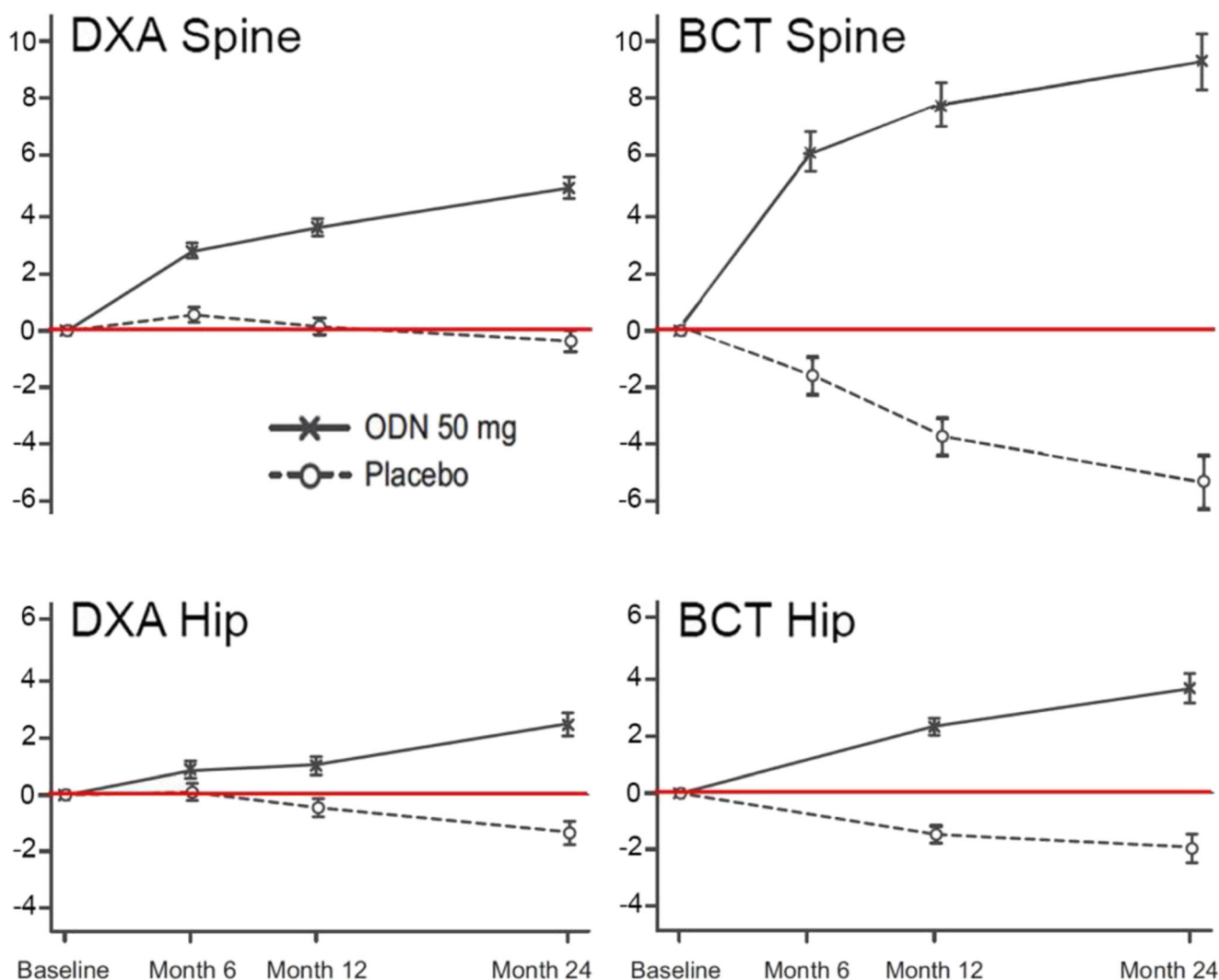
**Fig. 5** Percent change (mean ± SE) from baseline, 6–24 months, in BMD by DXA and bone strength by BCT at the spine and hip in postmenopausal women treated with odanacatib (X) or placebo (O). Adapted from Brixen [87], with axes scaled for uniformity

Those findings suggest that treatment-induced responses in BMD and strength can sometimes differ. Finally, the monkey study of denosumab discussed above showed that the magnitude of treatment-induced change in directly measured vertebral strength was correctly captured by BCT but not by overall BMD or bone mineral content [67]. Thus, in the context of fracture risk, BCT-measured strength responses may be easier to interpret than BMD responses.

## Opportunistic BCT may be cost-effective compared to usual-care DXA screening

To quantify the potential clinical impact of opportunistic BCT for the CT-patient, a theoretical cost-effectiveness study [94] simulated a hypothetical cohort of 1000 Medicare patients who had undergone an abdominal CT for any medical indication but were without a recent DXA. All those "DXA-lacking" CT patients were offered BCT, 90% were successfully tested,

and 50% of those who tested positive took generic alendronate for 2 years; hip fractures were then tracked over 5 years. That group was compared to usual care treatment for those patients: instead of being offered BCT, these hypothetical patients were tested annually with DXA at typical Medicare testing rates (9.5% for women, 1.7% for men [4]); and, like BCT, 50% of positive-testing patients were then treated with alendronate for 2 years. For both groups, over the 5-year observation period, patients were not monitored with any additional tests. Sensitivity and specificity for BCT and DXA were taken from the FOCUS real-world study [20].

For women, the results indicated that BCT testing would prevent over twice as many hip fractures and increase quality adjusted life years (QALY) almost fourfold, compared to managing these patients instead by usual-care DXA testing. The absolute benefit was greater in women but the relative benefit was greater in men (Fig. 6). The clinical efficacy increased further if more than 50% of positive-testing patients went on
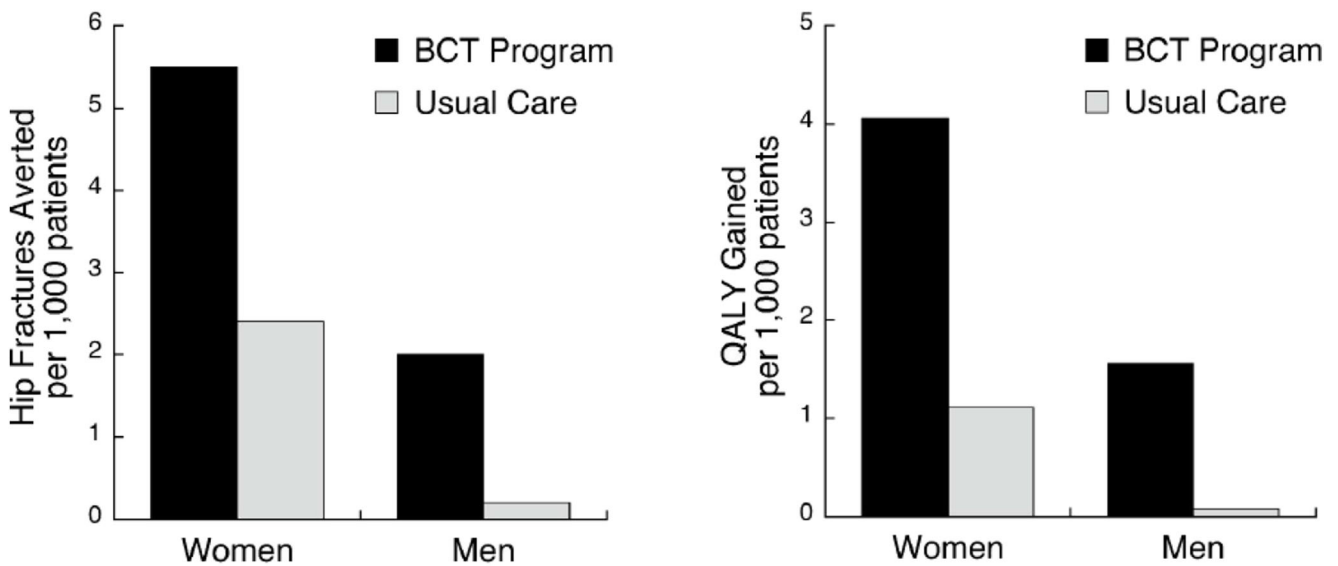
Fig. 6 Clinical utility of BCT and usual-care DXA testing versus no care. From Pisu [94]

treatment, if treatment was extended beyond 2 years, or if the BCT program was restricted to higher risk (e.g., older) patients. Likewise, clinical efficacy decreased if fewer than 50% of positive-testing patients went on treatment or if treatment only lasted for 1 year. These findings suggest that BCT, if widely implemented within a healthcare setting, could be cost-effective compared to usual-care DXA testing and in some circumstances may be cost saving. However, different healthcare systems and countries can have different clinical guidelines and payment and coverage policies for osteoporosis screening, which would be an additional issue to consider in assessing potential impact, cost-effectiveness, and feasibility of opportunistic BCT. Other groups have reported that BCT can also be cost-effective in certain circumstances even when BCT is used non-opportunistically with a dedicated CT scan [25, 95]. Despite the predictions of all these theoretical analyses, until implemented in practice in different settings, true efficacy and cost-effectiveness for BCT cannot be determined with certainty.

## Summary and discussion

Diagnostic screening for osteoporosis is critical for primary fracture prevention since it is the gateway to treatment [8, 46]. Osteoporosis is a condition of compromised bone strength leading to increased risk of fracture. Long and widely used for research purposes [14, 16, 35, 50, 71, 72, 96–110], BCT— aka finite element analysis of clinical-resolution CT scans—is well established for providing non-invasive measurements of bone strength [23–26]. Clinically, BCT is now a Medicare screening benefit in the USA, providing diagnostic-quality measurements of both bone strength and BMD from CT scans. When used opportunistically, and unlike simpler opportunistic analyses

of previously taken CT scans [111], the measurements from BCT can be used to identify osteoporosis and assess fracture risk and do not need confirmation by DXA. In that context, we reviewed the clinical evidence supporting opportunistic BCT, focusing on the VirtuOst implementation of BCT because it is the only FDA-cleared implementation of BCT available for patient care in the USA. With a clinical focus in mind, the body of evidence reviewed here supports the following:

- The BMD measurements from BCT can be used to identify osteoporosis and assess fracture risk using traditional clinical guidelines and FRAX or other risk calculators.
- Bone strength by BCT accurately measures bone strength in human cadavers and has accurately quantified treatment effects on bone strength in monkeys.
- Bone strength by BCT is associated with risk of hip, spine, and major osteoporotic fractures at least as strongly as by BMD by either DXA or quantitative CT, for both sexes.
- Bone strength by BCT is associated with fracture risk independent of BMD by DXA.
- Older patients classified by BCT as having either BMD-defined osteoporosis or fragile bone strength are at high risk of fracture.
- BCT measurements are precise and clinically reproducible.
- Although BCT can be used to monitor for a treatment response when the CT scan is obtained specifically for bone strength purposes, the major current clinical application of BCT is likely its opportunistic use. In this setting, monitoring for a treatment response by BCT is not currently recommended.
- Opportunistic BCT may be cost-effective compared to usual-care DXA screening.

When utilizing hip-containing CT exams of the abdomen or pelvis, BCT performed at the hip is particularly effective for identifying patients at high risk of hip fracture. With about seven million CT exams of the abdomen or pelvis being performed each year in the US Medicare population [27], opportunistic BCT at the hip should help improve osteoporosis care for this population and may do so in a cost-effective manner compared to usual care DXA testing [94]. For non-CT patients, DXA testing remains appropriate as per current guidelines, unless there are specific reasons to order a BCT with a dedicated CT—patients presenting without BMD-defined osteoporosis but with an otherwise unexplained fracture or patients with appreciable vertebral degeneration [55] or aortic calcification [47]. Others have reported that BCT may also be cost-effective in some situations when used for screening with a dedicated CT [25, 95].

A key issue in using BCT to manage patients is the validity and robustness of the interventional thresholds used for clinical decision-making. As reviewed here, the BCT interventional thresholds for both BMD and bone strength as measured by VirtuOst have been prospectively validated in a number of studies. Those studies were performed in the USA in racially [20] and geographically [16] diverse populations, in Iceland for a mostly Caucasian population [35], and for both sexes. Interestingly, preliminary data suggest that the bone strength thresholds used so far for BCT in the USA and Europe may also apply to native Koreans—because the relationship between BMD and bone strength appears to be uniform between Koreans and Caucasians [112], although fracture outcomes in these populations have not yet been assessed using these thresholds. In terms of robustness, the centralized laboratory delivery model for BCT—which can reach patients remotely—is designed to ensure quality and consistency of the BCT results, much like centralized laboratory services for blood and genetic tests. Thus, the BCT results via that delivery model should be sufficiently consistent and robust for clinical decision-making regardless of geographic location.

Despite the patient-convenience of opportunistic BCT, there are some logistical challenges for its current delivery model via a laboratory-based service. The associated clinical workflow is unique and may be challenging to widely implement, at least initially. For example, while BCT is now clinically available in the USA, its availability must meet patient privacy and HIPAA requirements and is currently limited to early-adopter healthcare providers. Internationally, there may be legal or patient-privacy barriers that prevent transmission of CT scans across national borders, perhaps requiring country-specific implementation. It is possible in the future that BCT becomes available for individual healthcare providers to perform locally without the need to transfer CT scans; for that scenario, issues of robustness, quality control, and standards may become more relevant. At present, while

the test is now covered and reimbursed by Medicare in the USA, due to its novelty, it will take time for healthcare providers to integrate BCT into its workflow and reimbursement payment pathways and for private insurers to set policy and payment rates for the test.

When using BCT clinically, physicians will occasionally encounter a patient who does not have BMD-defined osteoporosis but has fragile bone strength—are these patients good candidates for treatment? For example, Fig. 7 shows a subset of women and men in the AGES study from Iceland who were at high risk of a new hip fracture due to having fragile bone strength at the hip but without having BMD-defined osteoporosis at the femoral neck. Using just hip BCT measurements, the data reviewed here from the FOCUS and AGES studies established that women who test positive for fragile bone strength are at as high a risk of fracture as are those who test positive for BMD-defined osteoporosis by hip DXA, and
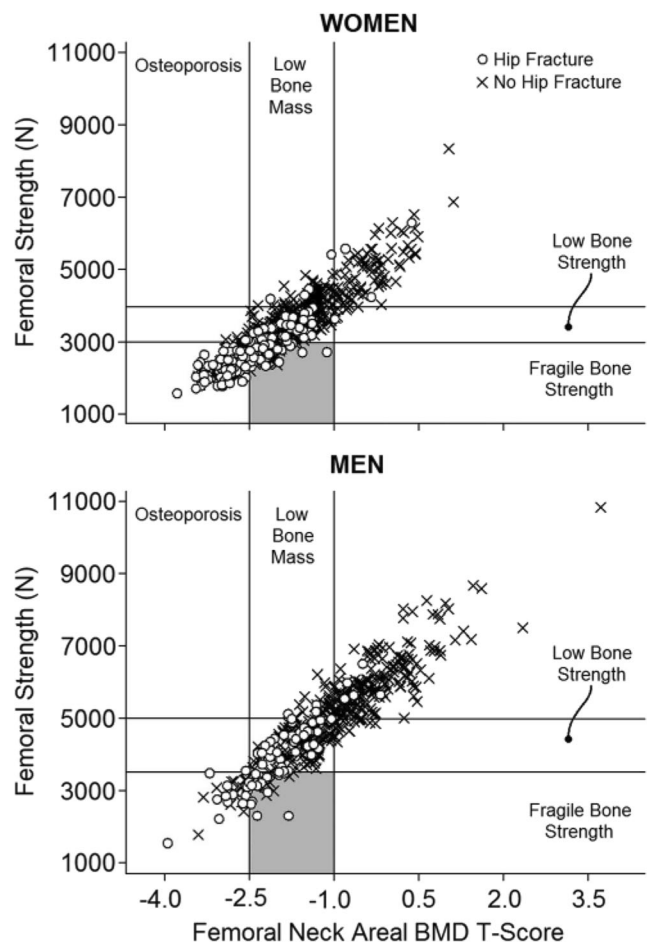


Fig. 7 Example of older women (top) and men (bottom) at high risk of fracture due to fragile bone strength but without BMD-defined osteoporosis at the femoral neck. The gray box contains all new hip fracture cases (open circles) who had fragile bone strength without BMD-defined osteoporosis at the femoral neck; plots taken directly from the AGES study of incident hip fracture in women and men in Iceland [35]. Note, few individuals in this study had BMD-defined osteoporosis at the femoral neck without fragile bone strength

FOCUS further established that these patients have about the same prevalence but are at higher risk of fracture than those testing positive by traditional hip/spine DXA. The issue then arises as to whether or not to treat such patients. As noted above, true osteoporosis is defined as "a skeletal disorder characterized by compromised *bone strength*, predisposing to *an increased risk of fracture*" [1, 2]. By that definition, patients with fragile bone strength have osteoporosis and should be considered for treatment. Consistent with this concept, current clinical practices already embrace identifying high-risk patients who present with BMD-defined osteopenia but who have other signs of compromised bone strength that place them at high risk of fracture. For example, one or more existing fractures in the absence of BMD-defined osteoporosis characterizes clinical osteoporosis because such fractures represent evidence of compromised bone strength—indeed, prevalent vertebral fractures are more highly associated with low bone strength by BCT than with low BMD by DXA [62–64].

Treating patients who have clinical signs of true osteoporosis, and not just those with BMD-defined osteoporosis, is effective. For example, drug trials to prevent new fractures demonstrate efficacy in BMD-defined osteopenic patients who have one or more existing vertebral fractures [113]. Also, alendronate has been shown to be more effective when given to patients who have lower BMD T-scores [114]; presumably, those patients on average will have lower bone strength because BMD T-scores and bone strength are correlated (see Fig. 7). Collectively, these findings indicate that osteoporosis medications are effective when given to patients who have (true) osteoporosis, namely, those who show evidence of compromised bone strength and increased risk of fracture. Realizing that individual treatment decisions always depend on patient-specific factors, older patients with fragile bone strength but without BMD-defined osteoporosis should be considered as candidates for treatment.

This clinical interpretation of BCT results is consistent with, and can easily be incorporated into, the latest US-based practice guidelines for assessing fracture risk and deciding who to treat. Traditionally with DXA, patients are classified for fracture risk on the basis of the BMD-based classifications listed in Table 1: patients with osteoporosis by BMD criteria are considered at high risk of fracture and thus are candidates for treatment, patients with low bone mass are considered at increased (or moderate) risk of fracture, and patients with normal bone mass are not at increased risk of fracture. That approach is the basis for the overall fracture risk classification from the BCT test. Practice guidelines in the USA also now include other metrics beyond BMD that classify who is at high risk of fracture [46, 79]. One important metric is the FRAX-calculated 10-year absolute risk of hip or other major osteoporotic fracture. That absolute risk in turn is often used with cut-points to classify high-risk patients for treatment: a 10-year hip fracture risk of 3% or more or a 10-year risk of major osteoporotic fractures of 20% or more [46, 79]. The other important metric is a history of previous fractures, particularly at the hip or spine. A new "very high-risk" category has been defined as the patient presenting with BMD-defined osteoporosis who also has multiple vertebral fractures [79]. The classifications from BCT, shown in Table 1, can be easily incorporated into these types of practice guidelines by expanding the definition of high-risk and very high-risk patients based on BMD criteria (T-score $\leq -2.5$ at the hip or spine) to also consider the presence of fragile bone strength. Further, while spinal coverage in abdominal CT scans typically only extends over the lumbar region, sagittal reconstructions from those scans can be useful nonetheless to opportunistically identify any prevalent vertebral fracture in that region [48, 49].

Another issue in using BCT clinically for patient care is what to do about monitoring for a treatment response once a BCT patient is placed on treatment. BCT, when used with a dedicated CT scan, can be used to monitor. However, given that *opportunistic* BCT has not yet been evaluated for monitoring treatment responses, we recommend that opportunistic BCT patients can be monitored for a treatment response by taking a baseline DXA once the patient is placed on treatment and then a follow-up monitoring DXA as appropriate. For repeated screening and fracture risk re-evaluation over time—which is distinct from monitoring a treatment response—any approved osteoporosis test, including opportunistic or dedicated BCT, can be used at any time point, and results interpreted appropriately.

BCT has some inherent limitations that should be recognized. One important clinical caveat, which applies also to any BMD measurement, is that BCT cannot detect any molecular level defects in bone that might compromise bone strength independently of the level of BMD. This is because clinical CT scans do not have molecular level resolution, and the computational finite element model used for BCT, which utilizes empirical relations from cadaver testing between BMD and the mechanical properties of trabecular and cortical bone, implicitly assumes a typical relation between CT-measured BMD and bone strength *at the tissue level*. Thus, physicians should be aware of this feature when interpreting results from a BCT test. For example, if a patient fractures without trauma and has both high BMD and high bone strength from BCT, one can reasonably interpret that finding as indirect but good evidence of compromised bone *quality* at the molecular level that presumably weakens the bone [115].

Another limitation of both BCT and BMD measurements for fracture risk assessment is that they only evaluate factors directly related to identifying the presence of osteoporosis; other factors related to fracture risk (e.g., fall risk) are not currently included. As a new modality for osteoporosis testing, and given its biomechanical and mechanistic nature, BCT is likely to evolve. As summarized elsewhere [24, 26], further improvements to BCT for measuring bone strength clinically

might result from incorporating such characteristics as trabecular orientation, enhanced cortical modeling, effects of aging on bone tissue ductility, and multiple sideway fall loading conditions—all topics of ongoing research [42, 97, 99, 100, 116]. In addition, beyond bone strength, other biomechanical factors are also important in fracture etiology: for example, fall risk, fall biomechanics, probabilistic loading, muscle strength, overall skeletal geometry, and soft tissue-related energy absorption and force attenuation [41]. These and other factors may eventually be integrated into BCT to more completely and mechanistically assess overall fracture risk.

Finally, it is noteworthy also that the BCT results presented here are limited to the specific implementation of BCT reviewed here (VirtuOst). Due to BCT's use to date primarily in research settings, different implementations of BCT by different research groups can provide different values of bone strength depending on a number of technical factors. Such factors include specifics of the image processing and calibration, the finite element modeling choices (e.g., type of finite element, material property mapping), and perhaps most importantly, how the virtual stress testing itself is configured, e.g., the bone orientation during the sideways fall or even the use of stance-like loading [26, 117]. Thus, all BCT is not the same and different implementations of BCT by different groups may produce different values of strength, requiring, for example, different classification cut-points. That said, one study reported different absolute measurements from two different implementations of BCT but similar measurements of treatment-induced percent changes [118], suggesting that BCT is more robust across different implementations for assessing temporal changes than absolute values. Related, the clinical term "BCT" is intended only to apply to finite element analysis of *clinical-resolution CT scans* and does not apply to finite element analysis of other types of medical images. For example, finite element analysis can be used with DXA [119, 120] and high-resolution peripheral quantitative CT [121, 122] scans, but both these applications are used primarily in research settings and are not available for clinical diagnostic purposes.

In summary, long used as a research tool, at this juncture, BCT is well established in terms of clinical efficacy and is now available clinically as a Medicare-reimbursed osteoporosis test in the USA. While it can be used with a dedicated CT, BCT is particularly well suited for opportunistic use in the patient who is already undergoing a hip- or spine-containing CT for other purposes. The data presented here indicate that BCT used in this way should provide a convenient, safe, and effective alternative to DXA for those patients not already undergoing DXA and who would benefit from a comprehensive osteoporosis assessment.

## Compliance with ethical standards

## References

1. U.S. Department of Health and Human Services (2004) Bone health and osteoporosis: a report of the surgeon general. U.S. Department of Health and Human Services, Office of the Surgeon General, Rockville

2. NIH Consensus Development Panel (2001) Osteoporosis prevention, diagnosis, and therapy. JAMA 285:785–795

3. International Osteoporosis Foundation (2020) Osteoporosis. http://wwwiofbonehealthorg/osteoporosis Accessed Feb 20 2020

4. Zhang J, Delzell E, Zhao H, Laster AJ, Saag KG, Kilgore ML, Morrisey MA, Wright NC, Yun H, Curtis JR (2012) Central DXA utilization shifts from office-based to hospital-based settings among Medicare beneficiaries in the wake of reimbursement changes. J Bone Miner Res 27:858–864

5. Lewiecki EM, Adler RA, Curtis JR (2016) Hip fractures and declining DXA testing: at a breaking point? J Bone Miner Res 31: S26

6. King AB, Fiorentino DM (2011) Medicare payment cuts for osteoporosis testing reduced use despite tests' benefit in reducing fractures. Health Aff (Millwood) 30:2362–2370

7. Lim SY, Lim JH, Nguyen D, Okamura R, Amiri HM, Calmes M, Nugent K (2013) Screening for osteoporosis in men aged 70 years and older in a primary care setting in the United States. Am J Mens Health 7:350–354

8. Siris ES, Modi A, Tang J, Gandhi S, Sen S (2014) Substantial under-treatment among women diagnosed with osteoporosis in a US managed-care population: a retrospective analysis. Curr Med Res Opin 30:123–130

9. Lewiecki E, Wright NC, Curtis JR, Siris E, Gagel RF, Saag KG, Singer AJ, Steven PM, Adler RA (2018) Hip fracture trends in the United States, 2002 to 2015. Osteoporos Int 29:717–722

10. Schousboe JT, Taylor BC, Fink HA, Kane RL, Cummings SR, Orwoll ES, Melton LJ 3rd, Bauer DC, Ensrud KE (2007) Cost-effectiveness of bone densitometry followed by treatment of osteoporosis in older men. JAMA 298:629–637

11. Nayak S, Greenspan SL (2016) Cost-effectiveness of osteoporosis screening strategies for men. J Bone Miner Res 31:1189–1199

12. Nayak S, Greenspan SL (2018) How can we improve osteoporosis care? A systematic review and meta-analysis of the efficacy of quality improvement strategies for osteoporosis. J Bone Miner Res 33:1585–1594

13. Wright NC, Looker AC, Saag KG, Curtis JR, Delzell ES, Randall S, Dawson-Hughes B (2014) The recent prevalence of osteoporosis and low bone mass in the United States based on bone mineral density at the femoral neck or lumbar spine. J Bone Miner Res 29: 2520–2526

14. Johannesdottir F, Thrall E, Muller J, Keaveny TM, Kopperdahl DL, Bouxsein ML (2017) Comparison of non-invasive assessments of strength of the proximal femur. Bone 105:93–102

15. Schuit SC, van der Klift M, Weel AE, de Laet CE, Burger H, Seeman E, Hofman A, Uitterlinden AG, van Leeuwen JP, Pols HA (2004) Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam study. Bone 34: 195–202

16. Orwoll ES, Marshall LM, Nielson CM, Cummings SR, Lapidus J, Cauley JA, Ensrud K, Lane N, Hoffmann PR, Kopperdahl DL, Keaveny TM (2009) Finite element analysis of the proximal femur and hip fracture risk in older men. J Bone Miner Res 24:475–483

17. Wainwright SA, Marshall LM, Ensrud KE, Cauley JA, Black DM, Hillier TA, Hochberg MC, Vogt MT, Orwoll ES (2005) Hip fracture in women without osteoporosis. J Clin Endocrinol Metab 90: 2787–2793

18. Oden A, McCloskey EV, Johansson H, Kanis JA (2013) Assessing the impact of osteoporosis on the burden of hip fractures. Calcif Tissue Int 92:42–49

19. Jiang X, Gruner M, Tremollieres F, Pluskiewicz W, Sornay-Rendu E, Adamczyk P, Schnatz PF (2017) Diagnostic accuracy of FRAX in predicting the 10-year risk of osteoporotic fractures using the USA treatment thresholds: a systematic review and meta-analysis. Bone 99:20–25

20. Adams AL, Fischer H, Kopperdahl DL, Lee DC, Black DM, Bouxsein ML, Fatemi S, Khosla S, Orwoll ES, Siris ES, Keaveny TM (2018) Osteoporosis and hip fracture risk from routine computed tomography scans: the Fracture, Osteoporosis, and CT Utilization Study (FOCUS). J Bone Miner Res 33:1291–1301

21. Keaveny TM (2010) Biomechanical computed tomography-noninvasive bone strength analysis using clinical computed tomography scans. Ann N Y Acad Sci 1192:57–65

22. Faulkner KG, Cann CE, Hasegawa BH (1991) Effect of bone distribution on vertebral strength: assessment with patient-specific nonlinear finite element analysis. Radiology 179:669–674

23. Zysset P, Qin L, Lang T, Khosla S, Leslie WD, Shepherd JA, Schousboe JT, Engelke K (2015) Clinical use of quantitative computed tomography-based finite element analysis of the hip and spine in the management of osteoporosis in adults: the 2015 ISCD official positions-part II. J Clin Densitom 18:359–392

24. Johannesdottir F, Allaire B, Bouxsein ML (2018) Fracture prediction by computed tomography and finite element analysis: current and future perspectives. Curr Osteoporos Rep 16:411–422

25. Viceconti M, Qasim M, Bhattacharya P, Li X (2018) Are CT-based finite element model predictions of femoral bone strengthening clinically useful? Curr Osteoporos Rep 16:216–223

26. Bouxsein ML, Zysset P, Gluer CC, McClung M, Biver E, Pierroz DD, Ferrari SL (2020) Perspectives on the non-invasive evaluation of femoral strength in the assessment of hip fracture risk. Osteoporos Int

27. The Centers for Medicare & Medicaid Services (2018) Part B national summary data file, Code Range: 03) Radiology (70000 - 79999) http://wwwcmsgov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Part-B-National-Summary-Data-File/Overviewhtml Accessed Feb 20 2020

28. FRAX (2019) Fracture risk assessment tool. http://wwwshefacuk/FRAX Accessed Feb 20 2020

29. American College of Radiology (2013) ACR practice guideline for the performance of quantitative computed tomography (QCT) bone densitometry (resolution 32). American College of Radiology, Reston

30. Engelke K, Adams JE, Armbrecht G, Augat P, Bogado CE, Bouxsein ML, Felsenberg D, Ito M, Prevrhal S, Hans DB, Lewiecki EM (2008) Clinical use of quantitative computed tomography and peripheral quantitative computed tomography in the management of osteoporosis in adults: the 2007 ISCD official positions. J Clin Densitom 11:123–162

31. Keaveny TM, McClung MR, Genant HK, Zanchetta JR, Kendler D, Brown JP, Goemaere S, Recknor C, Brandi ML, Eastell R, Kopperdahl DL, Engelke K, Fuerst T, Radcliffe HS, Libanati C (2014) Femoral and vertebral strength improvements in postmenopausal women with osteoporosis treated with denosumab. J Bone Miner Res 29:158–165

32. Lee DC, Hoffmann PF, Kopperdahl DL, Keaveny TM (2017) Phantomless calibration of CT scans for measurement of BMD and bone strength — inter-operator reanalysis precision. Bone 103:325–333

33. Wang X, Sanyal A, Cawthon PM, Palermo L, Jekir M, Christensen J, Ensrud KE, Cummings SR, Orwoll E, Black DM, Keaveny TM (2012) Prediction of new clinical vertebral fractures in elderly men using finite element analysis of CT scans. J Bone Miner Res 27:808–816

34. Allaire BT, Lu D, Johannesdottir F, Kopperdahl D, Keaveny TM, Jarraya M, Guermazi A, Bredella MA, Samelson EJ, Kiel DP, Anderson DE, Demissie S, Bouxsein ML (2019) Prediction of incident vertebral fracture using CT-based finite element analysis. Osteoporos Int 30:323–331

35. Kopperdahl DL, Aspelund T, Hoffmann PF, Sigurdsson S, Siggeirsdottir K, Harris TB, Gudnason V, Keaveny TM (2014) Assessment of incident spine and hip fractures in women and men using finite element analysis of CT scans. J Bone Miner Res 29: 570–580

36. Kopperdahl DL, Morgan EF, Keaveny TM (2002) Quantitative computed tomography estimates of the mechanical properties of human vertebral trabecular bone. J Orthop Res 20:801–805

37. Morgan EF, Keaveny TM (2001) Dependence of yield strain of human trabecular bone on anatomic site. J Biomech 34:569–577

38. Morgan EF, Bayraktar HH, Keaveny TM (2003) Trabecular bone modulus-density relationships depend on anatomic site. J Biomech 36:897–904

39. Bayraktar HH, Morgan EF, Niebur GL, Morris GE, Wong EK, Keaveny TM (2004) Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. J Biomech 37:27–35

40. Altai Z, Qasim M, Li X, Viceconti M (2019) The effect of boundary and loading conditions on patient classification using finite element predicted risk of fracture. Clin Biomech (Bristol, Avon) 68:137–143

41. Bhattacharya P, Altai Z, Qasim M, Viceconti M (2019) A multiscale model to predict current absolute risk of femoral fracture in a postmenopausal population. Biomech Model Mechanobiol 18:301–318

42. Nishiyama KK, Ito M, Harada A, Boyd SK (2014) Classification of women with and without hip fracture based on quantitative computed tomography and finite element analysis. Osteoporos Int 25:619–626

43. Michalski AS, Edwards WB, Boyd SK (2019) The influence of reconstruction kernel on bone mineral and strength estimates using quantitative computed tomography and finite element analysis. J Clin Densitom 22:219–228

44. Bauer JS, Henning TD, Mueller D, Lu Y, Majumdar S, Link TM (2007) Volumetric quantitative CT of the spine and hip derived

from contrast-enhanced MDCT: conversion factors. Am J Roentgenol 188:1294–1301

45. Weber NK, Fidler JL, Keaveny TM, Clarke BL, Khosla S, Fletcher JG, Lee DC, Pardi DS, Loftus EV Jr, Kane SV, Barlow JM, Murthy NS, Becker BD, Bruining DH (2014) Validation of a CT-derived method for osteoporosis screening in IBD patients undergoing contrast-enhanced CT enterography. Am J Gastroenterol 109:401–408

46. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R (2014) Clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 25:2359–2381

47. Fidler JL, Murthy NS, Khosla S, Clarke BL, Bruining DH, Kopperdahl DL, Lee DC, Keaveny TM (2016) Comprehensive assessment of osteoporosis and bone fragility with CT colonography. Radiology 278:172–180

48. Bauer JS, Muller D, Ambekar A, Dobritz M, Matsuura M, Eckstein F, Rummeny EJ, Link TM (2006) Detection of osteoporotic vertebral fractures using multidetector CT. Osteoporos Int 17:608–615

49. Lee SJ, Binkley N, Lubner MG, Bruce RJ, Ziemlewicz TJ, Pickhardt PJ (2016) Opportunistic screening for osteoporosis using the sagittal reconstruction from routine abdominal CT for combined assessment of vertebral fractures and density. Osteoporos Int 27:1131–1136

50. Nawathe S, Akhlaghpour H, Bouxsein ML, Keaveny TM (2014) Microstructural failure mechanisms in the human proximal femur for sideways fall loading. J Bone Miner Res 29:507–515

51. Poole KE, Skingle L, Gee AH, Turmezei TD, Johannesdottir F, Blesic K, Rose C, Vindlacheruvu M, Donell S, Vaculik J, Dungl P, Horak M, Stepan JJ, Reeve J, Treece GM (2017) Focal osteoporosis defects play a key role in hip fracture. Bone 94:124–134

52. Bolotin HH (2007) DXA in vivo BMD methodology: an erroneous and misleading research and clinical gauge of bone mineral status, bone fragility, and bone remodelling. Bone 41:138–154

53. Yu EW, Thomas BJ, Brown JK, Finkelstein JS (2012) Simulated increases in body fat and errors in bone mineral density measurements by DXA and QCT. J Bone Miner Res 27:119–124

54. Maldonado FJ, Al Bawardy BF, Nehra AK, Lee YS, Bruining DH, Adkins MC, Keaveny TM, Johnson MP, Fidler JL, McCollough CH, Fletcher JG (2019) Findings of CT-derived bone strength assessment in inflammatory bowel disease patients undergoing CT enterography in clinical practice. Inflamm Bowel Dis 25:1072–1079

55. Burch S, Feldstein M, Hoffmann PF, Keaveny TM (2016) Prevalence of poor bone quality in women undergoing spinal fusion using biomechanical-CT analysis. Spine (Phila Pa 1976) 41:246–252

56. Schwaiger BJ, Kopperdahl DL, Nardo L, Facchetti L, Gersing AS, Neumann J, Lee KJ, Keaveny TM, Link TM (2017) Vertebral and femoral bone mineral density and bone strength in prostate cancer patients assessed in phantomless PET/CT examinations. Bone 101:62–69

57. Binkley N, Kiebzak GM, Lewiecki EM, Krueger D, Gangnon RE, Miller PD, Shepherd JA, Drezner MK (2005) Recalculation of the NHANES database SD improves T-score agreement and reduces osteoporosis prevalence. J Bone Miner Res 20:195–201

58. Khoo BC, Brown K, Cann C, Zhu K, Henzell S, Low V, Gustafsson S, Price RI, Prince RL (2009) Comparison of QCT-derived and DXA-derived areal bone mineral density and T scores. Osteoporos Int 20:1539–1545

59. Ziemlewicz TJ, Maciejewski A, Binkley N, Brett AD, Brown JK, Pickhardt PJ (2016) Opportunistic quantitative CT bone mineral density measurement at the proximal femur using routine contrast-enhanced scans: direct comparison with DXA in 355 adults. J Bone Miner Res 31:1835–1840

60. Lee SJ, Anderson PA, Pickhardt PJ (2017) Predicting future hip fractures on routine abdominal CT using opportunistic osteoporosis screening measures: a matched case-control study. Am J Roentgenol 209:395–402

61. Cann CE, Genant HK, Kolb FO, Ettinger B (1985) Quantitative computed tomography for prediction of vertebral fracture risk. Bone 6:1–7

62. Melton LJ, Riggs BL, Keaveny TM, Achenbach SJ, Hoffmann PF, Camp JJ, Rouleau PA, Bouxsein ML, Amin S, Atkinson EJ, Robb RA, Khosla S (2007) Structural determinants of vertebral fracture risk. J Bone Miner Res 22:1885–1892

63. Melton LJ 3rd, Riggs BL, Keaveny TM, Achenbach SJ, Kopperdahl D, Camp JJ, Rouleau PA, Amin S, Atkinson EJ, Robb RA, Therneau TM, Khosla S (2010) Relation of vertebral deformities to bone density, structure, and strength. J Bone Miner Res 25:1922–1930

64. Anderson DE, Demissie S, Allaire BT, Bruno AG, Kopperdahl DL, Keaveny TM, Kiel DP, Bouxsein ML (2014) The associations between QCT-based vertebral bone measurements and prevalent vertebral fractures depend on the spinal locations of both bone measurement and fracture. Osteoporos Int 25:559–566

65. Black DM, Bouxsein ML, Marshall LM, Cummings SR, Lang TF, Cauley JA, Ensrud KE, Nielson CM, Orwoll ES (2008) Proximal femoral structure and the prediction of hip fracture in men: a large prospective study using QCT. J Bone Miner Res 23:1326–1333

66. Crawford RP, Cann CE, Keaveny TM (2003) Finite element models predict in vitro vertebral body compressive strength better than quantitative computed tomography. Bone 33:744–750

67. Lee DC, Varela A, Kostenuik PJ, Ominsky MS, Keaveny TM (2016) Finite element analysis of denosumab treatment effects on vertebral strength in ovariectomized cynomolgus monkeys. J Bone Miner Res 31:1586–1595

68. Buckley JM, Loo K, Motherway J (2007) Comparison of quantitative computed tomography-based measures in predicting vertebral compressive strength. Bone 40:767–774

69. Dall'Ara E, Schmidt R, Pahr D, Varga P, Chevalier Y, Patsch J, Kainberger F, Zysset P (2010) A nonlinear finite element model validation study based on a novel experimental technique for inducing anterior wedge-shape fractures in human vertebral bodies in vitro. J Biomech 43:2374–2380

70. Zysset PK, Dall'ara E, Varga P, Pahr DH (2013) Finite element analysis for prediction of bone strength. Bonekey Rep 2:386

71. Keyak JH, Rossi SA, Jones KA, Skinner HB (1998) Prediction of femoral fracture load using automated finite element modeling. J Biomech 31:125–133

72. Cody DD, Gross GJ, Hou FJ, Spencer HJ, Goldstein SA, Fyhrie DP (1999) Femoral strength is better predicted by finite element models than QCT and DXA. J Biomech 32:1013–1020

73. Dragomir-Daescu D, Op Den Buijs J, McEligot S, Dai Y, Entwistle RC, Salas C, Melton LJ 3rd, Bennet KE, Khosla S, Amin S (2011) Robust QCT/FEA models of proximal femur stiffness and fracture load during a sideways fall on the hip. Ann Biomed Eng 39:742–755

74. Dall'Ara E, Luisier B, Schmidt R, Kainberger F, Zysset P, Pahr D (2013) A nonlinear QCT-based finite element model validation study for the human femur tested in two configurations in vitro. Bone 52:27–38

75. Jayakar RY, Cabal A, Szumiloski J, Sardesai S, Phillips EA, Laib A, Scott BB, Pickarski M, Duong LT, Winkelmann CT, McCracken PJ, Hargreaves R, Hangartner TN, Williams DS (2012) Evaluation of high-resolution peripheral quantitative computed tomography, finite element analysis and biomechanical testing in a pre-clinical model of osteoporosis: a study with odanacatib treatment in the ovariectomized adult rhesus monkey. Bone 50:1379–1388

76. Amin S, Kopperdahl DL, Melton LJ 3rd, Achenbach SJ, Therneau TM, Riggs BL, Keaveny TM, Khosla S (2011) Association of hip strength estimates by finite-element analysis with fractures in women and men. J Bone Miner Res 26:1593–1600

77. Kanis JA, Oden A, Johnell O, Johansson H, De Laet C, Brown J, Burckhardt P, Cooper C, Christiansen C, Cummings S, Eisman JA, Fujiwara S, Gluer C, Goltzman D, Hans D, Krieg MA, La Croix A, McCloskey E, Mellstrom D, Melton LJ 3rd, Pols H, Reeve J, Sanders K, Schott AM, Silman A, Torgerson D, van Staa T, Watts NB, Yoshimura N (2007) The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. Osteoporos Int 18: 1033–1046

78. Keaveny TM, Kopperdahl DL, Melton LJ 3rd, Hoffmann PF, Amin S, Riggs BL, Khosla S (2010) Age-dependence of femoral strength in white women and men. J Bone Miner Res 25:994–1001

79. Eastell R, Rosen CJ, Black DM, Cheung AM, Murad MH, Shoback D (2019) Pharmacological management of osteoporosis in postmenopausal women: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 104:1595–1622

80. Fan B, Lu Y, Genant H, Fuerst T, Shepherd J (2010) Does standardized BMD still remove differences between Hologic and GE-lunar state-of-the-art DXA systems? Osteoporos Int 21:1227–1236

81. Cummings SR, Palermo L, Browner W, Marcus R, Wallace R, Pearson J, Blackwell T, Eckert S, Black D (2000) Monitoring osteoporosis therapy with bone densitometry: misleading changes and regression to the mean. JAMA 283:1318–1321

82. Keaveny TM, Donley DW, Hoffmann PF, Mitlak BH, Glass EV, San Martin JA (2007) Effects of teriparatide and alendronate on vertebral strength as assessed by finite element modeling of QCT scans in women with osteoporosis. J Bone Miner Res 22:149–157

83. Mawatari T, Miura H, Hamai S, Shuto T, Nakashima Y, Okazaki K, Kinukawa N, Sakai S, Hoffmann PF, Iwamoto Y, Keaveny TM (2008) Vertebral strength changes in rheumatoid arthritis patients treated with alendronate, as assessed by finite element analysis of clinical computed tomography scans: a prospective randomized clinical trial. Arthritis Rheum 58:3340–3349

84. Keaveny TM, Hoffmann PF, Singh M, Palermo L, Bilezikian JP, Greenspan SL, Black DM (2008) Femoral bone strength and its relation to cortical and trabecular changes after treatment with PTH, alendronate, and their combination as assessed by finite element analysis of quantitative CT scans. J Bone Miner Res 23: 1974–1982

85. Lewiecki EM, Keaveny TM, Kopperdahl DL, Genant HK, Engelke K, Fuerst T, Kivitz A, Davies RY, Fitzpatrick LA (2009) Once-monthly oral Ibandronate improves biomechanical determinants of bone strength in women with postmenopausal osteoporosis. J Clin Endocrinol Metab 94:171–180

86. Keaveny TM, McClung MR, Wan X, Kopperdahl DL, Mitlak BH, Krohn K (2012) Femoral strength in osteoporotic women treated with teriparatide or alendronate. Bone 50:165–170

87. Brixen K, Chapurlat R, Cheung AM, Keaveny TM, Fuerst T, Engelke K, Recker R, Dardzinski B, Verbruggen N, Ather S, Rosenberg E, de Papp AE (2013) Bone density, turnover, and estimated strength in postmenopausal women treated with odanacatib: a randomized trial. J Clin Endocrinol Metab 98: 571–580

88. Cosman F, Keaveny TM, Kopperdahl D, Wermers RA, Wan X, Krohn KD, Krege JH (2013) Hip and spine strength effects of adding versus switching to teriparatide in postmenopausal women with osteoporosis treated with prior alendronate or raloxifene. J Bone Miner Res 28:1328–1336

89. Kleerekoper M, Greenspan SL, Lewiecki EM, Miller PD, Kendler DL, Maricic M, Keaveny TM, Kopperdahl DL, Ruff VA, Wan X,

Janos B, Krohn K (2014) Assessing the effects of teriparatide treatment on bone mineral density, bone microarchitecture, and bone strength. J Bone Joint Surg Am 96:e90

90. Snyder PJ, Kopperdahl DL, Stephens-Shields AJ, Ellenberg SS, Cauley JA, Ensrud KE, Lewis CE, Barrett-Connor E, Schwartz AV, Lee DC, Bhasin S, Cunningham GR, Gill TM, Matsumoto AM, Swerdloff RS, Basaria S, Diem SJ, Wang C, Hou X, Cifelli D, Dougar D, Zeldow B, Bauer DC, Keaveny TM (2017) Effect of testosterone treatment on volumetric bone density and strength in older men with low testosterone: a controlled clinical trial. JAMA Intern Med 177:471–479

91. Keaveny TM, Crittenden DB, Bolognese MA, Genant HK, Engelke K, Oliveri B, Brown JP, Langdahl BL, Yan C, Grauer A, Libanati C (2017) Greater gains in spine and hip strength for romosozumab compared to teriparatide in postmenopausal women with low bone mass. J Bone Miner Res 32:1956–1962

92. Langdahl BL, Libanati C, Crittenden DB, Bolognese MA, Brown JP, Daizadeh NS, Dokoupilova E, Engelke K, Finkelstein JS, Genant HK, Goemaere S, Hyldstrup L, Jodar-Gimeno E, Keaveny TM, Kendler D, Lakatos P, Maddox J, Malouf J, Massari FE, Molina JF, Ulla MR, Grauer A (2017) Romosozumab (sclerostin monoclonal antibody) versus teriparatide in postmenopausal women with osteoporosis transitioning from oral bisphosphonate therapy: a randomised, open-label, phase 3 trial. Lancet 390:1585–1594

93. Howe JG, Hill RS, Stroncek JD, Shaul JL, Favell D, Cheng RR, Engelke K, Genant HK, Lee DC, Keaveny TM, Bouxsein ML, Huber B (2019) Treatment of bone loss in proximal femurs of postmenopausal osteoporotic women with AGN1 local osteo-enhancement procedure (LOEP) increases hip bone mineral density and hip strength: a long-term prospective cohort study. Osteoporos Int

94. Pisu M, Kopperdahl DL, Lewis CE, Saag KG, Keaveny TM (2019) Cost-effectiveness of osteoporosis screening using biomechanical computed tomography for patients with a previous abdominal CT. J Bone Miner Res 34:1229–1239

95. Agten CA, Ramme AJ, Kang S, Honig S, Chang G (2017) Cost-effectiveness of virtual bone strength testing in osteoporosis screening programs for postmenopausal women in the United States. Radiology 285:506–517

96. Falcinelli C, Schileo E, Balistreri L, Baruffaldi F, Bordini B, Viceconti M, Albisinni U, Ceccarelli F, Milandri L, Toni A, Taddei F (2014) Multiple loading conditions analysis can improve the association between finite element bone strength estimates and proximal femur fractures: a preliminary study in elderly women. Bone 67:71–80

97. Schileo E, Dall'ara E, Taddei F, Malandrino A, Schotkamp T, Baleani M, Viceconti M (2008) An accurate estimation of bone density improves the accuracy of subject-specific finite element models. J Biomech 41:2483–2491

98. Grassi L, Schileo E, Taddei F, Zani L, Juszczyk M, Cristofolini L, Viceconti M (2012) Accuracy of finite element predictions in sideways load configurations for the proximal human femur. J Biomech 45:394–399

99. Enns-Bray WS, Bahaloo H, Fleps I, Ariza O, Gilchrist S, Widmer R, Guy P, Palsson H, Ferguson SJ, Cripton PA, Helgason B (2018) Material mapping strategy to improve the predicted response of the proximal femur to a sideways fall impact. J Mech Behav Biomed Mater 78:196–205

100. Bahaloo H, Enns-Bray WS, Fleps I, Ariza O, Gilchrist S, Soyka RW, Guy P, Palsson H, Ferguson SJ, Cripton PA, Helgason B (2018) On the failure initiation in the proximal human femur under simulated sideways fall. Ann Biomed Eng 46:270–283

101. Ariza O, Gilchrist S, Widmer RP, Guy P, Ferguson SJ, Cripton PA, Helgason B (2015) Comparison of explicit finite element and

mechanical simulation of the proximal femur during dynamic drop-tower testing. J Biomech 48:224–232

102. Bessho M, Ohnishi I, Matsumoto T, Ohashi S, Matsuyama J, Tobita K, Kaneko M, Nakamura K (2009) Prediction of proximal femur strength using a CT-based nonlinear finite element method: differences in predicted fracture load and site with changing load and boundary conditions. Bone 45:226–231

103. Haider IT, Speirs AD, Frei H (2013) Effect of boundary conditions, impact loading and hydraulic stiffening on femoral fracture strength. J Biomech 46:2115–2121

104. Keyak JH, Sigurdsson S, Karlsdottir GS, Oskarsdottir D, Sigmarsdottir A, Kornak J, Harris TB, Sigurdsson G, Jonsson BY, Siggeirsdottir K, Eiriksdottir G, Gudnason V, Lang TF (2013) Effect of finite element model loading condition on fracture risk assessment in men and women: the AGES-Reykjavik study. Bone 57:18–29

105. Kheirollahi H, Luo Y (2015) Identification of high stress and strain regions in proximal femur during single-leg stance and sideways fall using QCT-based finite element model. Int J Med Health Biomed Bioeng Pharm Eng 9:633–640

106. Lang TF, Sigurdsson S, Karlsdottir G, Oskarsdottir D, Sigmarsdottir A, Chengshi J, Kornak J, Harris TB, Sigurdsson G, Jonsson BY, Siggeirsdottir K, Eiriksdottir G, Gudnason V, Keyak JH (2012) Age-related loss of proximal femoral strength in elderly men and women: The Age Gene/Environment Susceptibility Study–Reykjavik. Bone 50:743–748

107. Lotz JC, Cheal EJ, Hayes WC (1995) Stress distributions within the proximal femur during gait and falls: implications for osteoporotic fracture. Osteoporos Int 5:252–261

108. Schileo E, Balistreri L, Grassi L, Cristofolini L, Taddei F (2014) To what extent can linear finite element models of human femora predict failure under stance and fall loading configurations? J Biomech 47:3531–3538

109. Schileo E, Taddei F, Cristofolini L, Viceconti M (2008) Subject-specific finite element models implementing a maximum principal strain criterion are able to estimate failure risk and fracture location on human femurs tested in vitro. J Biomech 41:356–367

110. Varga P, Schwiedrzik J, Zysset PK, Fliri-Hofmann L, Widmer D, Gueorguiev B, Blauth M, Windolf M (2016) Nonlinear quasi-static finite element simulations predict in vitro strength of human proximal femora assessed in a dynamic sideways fall setup. J Mech Behav Biomed Mater 57:116–127

111. Dagan N, Elnekave E, Barda N, Bregman-Amitai O, Bar A, Orlovsky M, Bachmat E, Balicer RD (2020) Automated opportunistic osteoporotic fracture risk assessment using computed tomography scans to aid in FRAX underutilization. Nat Med 26:77–82

112. Lee DC, Hong N, Khosla S, Rhee Y, Keaveny TM (2019) Prevalence of spinal osteoporosis in women and men considering both bone strength and volumetric BMD — a comparison of Caucasians (in the United States) and Koreans (in Korea). J Bone Miner Res 34:100

113. Bone HG, Dempster DW, Eisman JA, Greenspan SL, McClung MR, Nakamura T, Papapoulos S, Shih WJ, Rybak-Feiglin A, Santora AC, Verbruggen N, Leung AT, Lombardi A (2015) Odanacatib for the treatment of postmenopausal osteoporosis: development history and design and participant characteristics of LOFT, the Long-Term Odanacatib Fracture Trial. Osteoporos Int 26:699–712

114. Cummings SR, Black DM, Thompson DE, Applegate WB, Barrett-Connor E, Musliner TA, Palermo L, Prineas R, Rubin SM, Scott JC, Vogt T, Wallace R, Yates AJ, LaCroix AZ (1998) Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. JAMA 280:2077–2082

115. Tucci JR, Whitford GM, McAlister WH, Novack DV, Mumm S, Keaveny TM, Whyte MP (2017) Skeletal fluorosis due to inhalation abuse of a difluoroethane-containing computer cleaner. J Bone Miner Res 32:188–195

116. Enns-Bray WS, Bahaloo H, Fleps I, Pauchard Y, Taghizadeh E, Sigurdsson S, Aspelund T, Buchler P, Harris T, Gudnason V, Ferguson SJ, Palsson H, Helgason B (2019) Biofidelic finite element models for accurately classifying hip fracture in a retrospective clinical study of elderly women from the AGES Reykjavik cohort. Bone 120:25–37

117. Michalski AS, Amin S, Cheung AM, Cody DD, Keyak JH, Lang TF, Nicolella DP, Orwoll ES, Boyd SK, Sibonga JD (2019) Hip load capacity cut-points for astronaut skeletal health NASA finite element strength task group recommendations. NPJ Microgravity 5:6

118. Zysset P, Pahr D, Engelke K, Genant HK, McClung MR, Kendler DL, Recknor C, Kinzl M, Schwiedrzik J, Museyko O, Wang A, Libanati C (2015) Comparison of proximal femur and vertebral body strength improvements in the FREEDOM trial using an alternative finite element methodology. Bone 81:122–130

119. Terzini M, Aldieri A, Rinaudo L, Osella G, Audenino AL, Bignardi C (2019) Improving the hip fracture risk prediction through 2D finite element models from DXA images: validation against 3D models. Front Bioeng Biotechnol 7:220

120. Yang S, Leslie WD, Luo Y, Goertzen AL, Ahmed S, Ward LM, Delubac I, Lix LM (2018) Automated DXA-based finite element analysis for hip fracture risk stratification: a cross-sectional study. Osteoporos Int 29:191–200

121. Westbury LD, Shere C, Edwards MH, Cooper C, Dennison EM, Ward KA (2019) Cluster analysis of finite element analysis and bone microarchitectural parameters identifies phenotypes with high fracture risk. Calcif Tissue Int 105:252–262

122. Stein EM, Liu XS, Nickolas TL, Cohen A, Thomas V, McMahon DJ, Zhang C, Yin PT, Cosman F, Nieves J, Guo XE, Shane E (2010) Abnormal microarchitecture and reduced stiffness at the radius and tibia in postmenopausal women with fractures. J Bone Miner Res 25:2572–2581