Title

Osteoporosis and Hip Fracture Risk From Routine Computed Tomography Scans: The Fracture, Osteoporosis, and CT Utilization Study (FOCUS)

Journal

Journal of Bone and Mineral Research, 33(7)

ISSN

0884-0431

Authors

Adams, Annette L
Fischer, Heidi
Kopperdahl, David L
et al.

Publication Date

2018-07-01

DOI

10.1002/jbmr.3423

Peer reviewed

HHS Public Access

Author manuscript

J Bone Miner Res. Author manuscript; available in PMC 2019 July 01.

Published in final edited form as:

J Bone Miner Res. 2018 July ; 33(7): 1291–1301. doi:10.1002/jbmr.3423.

Osteoporosis and Hip Fracture Risk From Routine Computed Tomography Scans: The Fracture, Osteoporosis, and CT Utilization Study (FOCUS)

Annette L Adams[1], **Heidi Fischer**[1], **David L Kopperdahl**[2], **David C Lee**[2], **Dennis M Black**[3], **Mary L Bouxsein**[4], **Shireen Fatemi**[5], **Sundeep Khosla**[6], **Eric S Orwoll**[7], **Ethel S Siris**[8], and **Tony M Keaveny**[9]

[1]Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA, USA

[2]O.N. Diagnostics, Berkeley, CA, USA

[3]Departments of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA

[4]Orthopedic Biomechanics Laboratory, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA

[5]Department of Endocrinology, Kaiser Permanente Southern California, Panorama City, CA, USA

[6]Kogod Center on Aging and Division of Endocrinology, Mayo Clinic, Rochester, MN, USA

[7]Bone and Mineral Unit, Oregon Health and Science University, Portland, OR, USA

[8]Toni Stabile Osteoporosis Center, Department of Medicine, Columbia University Medical Center, New York, NY, USA

[9]Departments of Mechanical Engineering and Bioengineering, University of California, Berkeley, CA, USA

Abstract

Methods now exist for analyzing previously taken clinical computed tomography (CT) scans to measure a dual-energy X-ray absorptiometry (DXA)-equivalent bone mineral density (BMD) at the hip and a finite element analysis–derived femoral strength. We assessed the efficacy of this "biomechanical CT" (BCT) approach for identifying patients at high risk of incident hip fracture in a large clinical setting. Using a case-cohort design sampled from 111,694 women and men aged 65 or older who had a prior hip CT scan, a DXA within 3 years of the CT, and no prior hip fracture, we compared those with subsequent hip fracture (n = 1959) with randomly selected sex-

Address correspondence to: Annette L Adams, PhD, Department of Research and Evaluation, 100 S Los Robles, 2nd Floor, Pasadena, CA 91101, USA. Annette.L.Adams@kp.org.

Disclosures

ALA has received research support from Amgen and Merck for this and other unrelated projects. DLK is an employee of and has equity in O.N. Diagnostics, developer of VirtuOst. DCL is an employee of and has equity in O.N. Diagnostics. TMK has been a consultant for Amgen, AgNovos Healthcare, and O.N. Diagnostics and has equity in O.N. Diagnostics. TMK, DLK, and DCL have financial interests in O.N. Diagnostics and both they and the company may benefit from the results of this work. All other authors report no disclosures.

Additional Supporting Information may be found in the online version of this article.

stratified controls (n = 1979) and analyzed their CT scans blinded to all other data. We found that the age-, race-, and body mass index (BMI)-adjusted hazard ratio (HR; per standard deviation) for femoral strength was significant before (women: HR = 2.8, 95% confidence interval [CI] 2.2–3.5; men: 2.8, 2.1–3.7) and after adjusting also for the (lowest) hip BMD *T*-score by BCT (women: 2.1, 1.4–3.2; men: 2.7, 1.6–4.6). The hazard ratio for the hip BMD *T*-score was similar between BCT and DXA for both sexes (women: 2.1, 1.8–2.5 BCT versus 2.1, 1.7–2.5 DXA; men: 2.8, 2.1–3.8 BCT versus 2.5, 2.0–3.2 DXA) and was higher than for the (lowest) spine/hip BMD *T*-score by DXA (women: 1.6, 1.4–1.9; men: 2.1, 1.6–2.7). Compared with the latter as a clinical-practice reference and using both femoral strength and the hip BMD T-score from BCT, sensitivity for predicting hip fracture was higher for BCT (women: 0.66 versus 0.59; men: 0.56 versus 0.48), with comparable respective specificity (women: 0.66 versus 0.67; men: 0.76 versus 0.78). We conclude that BCT analysis of previously acquired routine abdominal or pelvic CT scans is at least as effective as DXA testing for identifying patients at high risk of hip fracture.

Keywords

COMPUTED TOMOGRAPHY; DXA; OSTEOPOROSIS; SCREENING; FRACTURE RISK ASSESSMENT; BONE STRENGTH

Introduction

Osteoporosis reduces bone mineral density (BMD) and bone strength and causes millions of fractures, with hip fractures being the most serious and costly.[1,2] Despite the availability of treatments that reduce fracture risk, osteoporosis remains substantially underdiagnosed.[3–5] The clinical standard for diagnosing osteoporosis is based on measuring BMD using dual-energy X-ray absorptiometry (DXA), but DXA screening rates remain low[6,7] and are on the order of 5.5% (sex-pooled) in the Medicare population.[8]

One effective option to address this care gap would be to expand options for osteoporosis screening while maintaining efficacy for predicting hip fracture. Toward that end, it has been proposed to reanalyze previously taken routine clinical computed tomography (CT) scans that contain the proximal femur to measure a DXA-equivalent hip BMD, without the need for using an external calibration phantom during patient imaging.[9–17]Because millions of hip-containing CT scans are taken annually,[18]the “add-on” nature of this ancillary approach offers the potential for widespread adoption. In a potential improvement, it is now also possible to perform a “biomechanical CT” (BCT)[19] analysis on such previously taken hip-containing CT scans.[14–16] The BCT analysis uses finite element analysis–based virtual stress testing, based on the information contained within the CT scan, to provide an estimate of the breaking strength of the femur and has been validated for predicting hip fracture in various research settings;[20–26] its clinical implementation also includes measurement of the DXA-equivalent hip BMD *T*-score.[14–16] Our goal in this study was to evaluate the diagnostic performance of ancillary BCT when used in this fashion in a large multicenter clinical setting. A secondary goal was to assess, for the BCT analysis, whether femoral strength can add clinical value to the hip BMD *T*-score in identifying patients at high risk of hip fracture; all results were compared against DXA as a clinical reference.

Materials and Methods

Study design

The Fracture, Osteoporosis, and CT Utilization Study (FOCUS) was a retrospectively performed case-cohort study using preexisting, anonymized clinical data from electronic health records to predict new hip fractures. The study was approved by the local Institutional Review Board at Kaiser Permanente Southern California (KPSC) with a waiver of informed consent.

Patient population

Our study cohort included KPSC members aged 65 years or older who had any abdominal or pelvic CT exam between January 1, 2006, and December 31, 2014; a DXA exam within 3 years of the CT exam; and no hip fracture before either the CT or DXA exam. Patients were excluded if they: 1) had a diagnosis of bone pathology within 3 months of the DXA or CT (ie, metastatic cancer, multiple myeloma, Paget's disease, osteogenesis imperfecta, hypophosphatasia, or hypo/hypercalcemia); 2) were of unknown race/ethnicity; or 3) experienced a high-energy hip fracture (based on ICD-9 codes).

Of the 728,089 patients who had a qualifying CT exam, 111,694 patients (64,992 women and 46,702 men) met the inclusion and exclusion criteria and comprised the source population. A case-cohort study was then conducted within the source population to examine time-to-event outcomes (incident hip fracture), with separate case-cohorts for the male and female populations (Supplemental Fig. S1). The overall sample size for analysis was limited by resources to approximately 4000 patients. Thus, we went back in time sufficiently far (to 2006) so as to be able to identify approximately 2000 cases and we compared those to approximately 2000 controls for a 1:1 case-to-control ratio, which tends to maximize power in case-cohort studies for a limited sample size.[27] Cases were identified as all patients in the source population who had a first fragility (non-traumatic) hip fracture during the study period. Subcohorts were then selected via sex-stratified random sampling of the source population (2.1% of the women and 1.4% of the men—about 3% higher than the approximate fracture rates in the source population) to provide an approximate 1:1 ratio of cases to controls for each sex. These subcohorts included any sampled hip-fracture cases, combined with all remaining cases in the source population to form the sex-specific case-cohorts.

All cohort members were followed from the earliest of their DXA or CT exams until they withdrew from KPSC, had a disqualifying diagnosis, or September 30, 2015 (end of study observation), whichever came first. The resulting sample comprised 3938 patients: 2690 women (1340 cases) and 1248 men (619 cases). Descriptive analyses confirmed that the analytic subcohorts and cases represented the source population for the various patient parameters, including age, race/ethnicity, body mass index (BMI), and BMD as measured by DXA.

CT scans

Qualifying CT exams were any abdominal or pelvic CT scans with or without contrast enhancement, identified by CPT code (72192–94, 74150/60/70, 74174–78). Of the 3938 patients selected for the case-cohort, 3861 patient-CT scans were successfully retrieved from the KPSC records and de-identified, and were sent to the core lab for BCT analysis. After excluding scans that were found to be erroneously included in the study (n = 205), either hip BMD and/or femoral strength was successfully measured by BCT for 86% (3160/3656) of the remaining scans (see Supplemental Materials for other excluding reasons). Those 3160 CT scans had been acquired in 14 different hospitals on 80 different CT scanners, representing 15 different scanner models and using 26 different types/versions of system software, most scans being taken on scanners manufactured by GE Medical Systems (99.7% = 3151/3160; five scans on Philips, two scans on Siemens, and two scans on Toshiba). Most scans were acquired at 120 kVp (97% = 3069/3160; nine at 100 kVp and 82 at 140 kVp), were archived as derived secondary compressed images (93% = 2950/3160) using a large field of view, used a standard kernel (98.8% = 3123/3160), and had a slice thickness in the range 0.5 to 5.0 mm (<1.0 mm, n = 250; 1 to 2 mm, n = 143; 2.5 mm, n = 978; 3 to 4 mm, n = 953; 5.0 mm, n = 836). Some type of intravenous contrast enhancement was used in 64% (2008/3160) of the scans.

BCT analysis

All BCT analyses were performed at a central core laboratory (O.N. Diagnostics, Berkeley, CA, USA), which was blinded to the hip fracture status, clinical factors, and DXA data (as were all investigators other than the KPSC personnel). The BCT analysis (VirtuOst 2.1, O.N. Diagnostics) was performed nominally on the left hip. As described in more detail elsewhere,[14,15,17,20,28] the BCT analysis uses a multistep process to provide an estimate of femoral strength for a simulated fall to the side of the hip (Fig. 1). First, the proximal femur is separated from any surrounding tissues and resampled into 1.5-mm-sized voxels in a standardized sideways-fall coordinate system. Second, independent of this process, a patient-specific phantomless calibration is performed, using the air and hip fat in the CT scan as reference materials.[14,15,17,20,28] Third, the calibrated BMD values of each voxel in the scan are converted into bone mechanical properties using empirical relations derived from human cadaver studies. Fourth, a finite element model is constructed by converting each voxel into an 8-noded cube-shaped finite element. And finally, via non-linear finite element analysis, virtual loads applied to the bone are gradually increased until the bone begins to break, thus providing an estimate of the breaking strength of the femur (defined as the force, in units of newtons, at an overall deformation of 4.0% strain). This approach has been validated against cadaver strength measurements.[29]

The VirtuOst-based BCT analysis used in FOCUS also provides DXA-equivalent BMD T-scores at the hip.[14,15,17,20,28] In two prior studies[14,15] that analyzed previously taken routine clinical CT scans in which the BCT measurements were blinded to the DXA data, the femoral neck BMD T-scores from VirtuOst agreed well with those from DXA. In the first study on CT enterography scans of 136 patients (women and men),[14] the scans containing intravenous contrast, there was a high correlation (R^2= 0.84) between VirtuOst and DXA and a $Y = X$ type of agreement (ie, the data fell close to a line with a slope of 1.0

and a zero intercept); there was also a small but significant paired difference (BCT lower) of 0.18 *T*-score units that was associated with BMI effects, higher BMI values significantly increasing the *T*-score for DXA compared with BCT. In the second study on CT colonography exams,[15] also on 136 patients (all women) and the scans containing oral contrast, there was an equally high correlation ($R^2 = 0.84$) between VirtuOst and DXA, with no significant difference in paired measurements between the two. In FOCUS, we defined the "hip BMD T-score" as the lower BMD *T*-score of the femoral neck and total hip regions. Following clinical guidelines,[30] *T*-scores were calculated from NHANESIII[31] using white young-reference values for all patients, but sex-specific values were used to increase sensitivity for the men. Using the hip BMD *T*-score and the femoral strength measurements, preestablished cut-points were then used to classify patients for "hip osteoporosis" (hip BMD *T*-score ≤–2.5, for both sexes) and "fragile bone strength" (femoral strength ≤3000 N for women or ≤3500 N for men).[20] The probability of incident hip fracture at these cut-points is similar for hip osteoporosis (by BCT) and fragile bone strength for both sexes.[20]

Some minor adjustments were made to the above process to accommodate the wide variation in some of the CT acquisition settings. Because both the strength and BMD measurements can be affected by slice thickness/increment of the CT reconstruction, particularly when it is above 3.0 mm, and because the CT attenuation of the femoral bone increases slightly when using intravenous contrast,[13,32] we made minor adjustments for both of these conditions (Supplemental Materials). Both adjustments were made blinded to all clinical data (including DXA).

Clinical factors and DXA

Data from the electronic health records at KPSC were obtained for the following clinical factors: age, sex, race/ethnicity, exclusion conditions (Paget's, metastatic cancer, etc.), BMI, and osteoporosis medication use. Race/ethnicity was characterized using four categories (Asian, black, Hispanic, and white; patients in other categories were classified as white). Using pharmacy data, we defined "recently untreated" patients as those who did not fill two or more prescriptions for an osteoporosis therapeutic agent within 1 year of either the CT or DXA exam. All DXA exams at KPSC were performed using Lunar (GE Healthcare, Madison, WI, USA) DXA scanners as part of the patient's medical care.[33] We report the hip BMD *T*-score for direct comparison with BCT and also the widely used[34] lowest BMD *T*-score across both the hip (femoral neck or total hip) and spine ("hip/spine" BMD *T*-score).

Statistical analyses

All statistical analyses were performed at KPSC using version 3.3.3 of the R software. Descriptive analyses examined the distributions (and frequencies) for the continuous (and categorical) variables. Clinical performance for all BMD *T*-scores and femoral strength was quantified using hazard ratios and 95% confidence intervals estimated from weighted Cox proportional regressions; these continuous variables were first standardized by subtracting the mean and dividing by the sex-pooled standard deviation of available data. All regressions used Breslow calibrated weights to account for the case-cohort design.[35] Follow-up began at the time of the DXA exam for DXA testing and the CT exam for BCT analysis. Age was

used as the underlying timescale to enable a more flexible association with baseline hazard.[36] All models were adjusted for race/ethnicity and BMI, and the femoral strength model was additionally adjusted by the hip BMD *T*-score (by BCT). To assess the level of risk of hip fracture for patients testing positive for hip osteoporosis (by BCT and DXA), hip/spine osteoporosis, fragile bone strength, and "high risk" (defined two ways: having either hip osteoporosis by BCT or fragile bone strength, or having both conditions), we used the same methods to also compute hazard ratios for these binary classifications.

In our main analysis, the hazard ratios could not be directly compared between DXA and BCT in the same regression model because of unequal follow-up times for each exam. Because hazard ratios are normally distributed on the log scale, we compared these hazard ratios using z tests for the difference between two statistical quantities in independent populations. These tests are conservative because the BCT and DXA measurements are positively correlated.

We also performed a number of stratified analyses. First, to reduce any confounding effects of osteoporosis treatment on the results, we performed analyses just on the recently untreated patients. For the subset of these patients who had 5-year follow-up, we calculated sensitivity, specificity, and AUC for predicting hip fracture within 5 years. For the BCT test, we assessed the individual contributions of hip osteoporosis (by BCT) and fragile bone strength to 5-year prediction of hip fracture by examining odds ratios (OR) and 95% confidence intervals (CI) in sex-stratified multivariate logistic regressions. We also assessed combining the osteoporosis and fragile bone strength classifications for identifying patients who would fracture, using both "or" and "and" logicals, the latter being more conservative. Second, we also ran stratified analyses to assess any confounding role of the main CT-imaging parameters. For this, we repeated the main hazard ratio analysis but separately in subcohorts of patients having scans above versus below 3.0-mm slice thickness and for patients having scans with versus without intravenous contrast enhancement. In addition to these stratified analyses, we used a two-factor analysis of covariance (BMI as the covariate) to test for any dependence on the difference in femoral neck BMD between BCT and DXA on the two imaging parameters (presence of intravenous contrast; thickness >3.0 mm); this analysis was restricted to women not on recent treatment and who had the CT and DXA exams within 6 months ($n = 536$), in order to minimize any confounding effects of treatment or time between exams on the comparison; we focused on femoral neck BMD because that measurement was available for most patients and is most directly comparable between DXA and BCT. Additional analyses (Supplemental Materials) were performed to confirm the equivalence of the BMD *T*-score as measured by BCT versus DXA.

Results

For both sexes, patients who fractured their hip were older and had lower BMD, lower BMI, and included more whites and fewer Asians, blacks, and Hispanics (Table 1). Median follow-up was similar between the sexes (3.6 versus 3.8 years for the women versus men) but was slightly shorter for the cases than controls (respectively: 3.2 versus 4.0 years women; 3.4 versus 4.4 years men). The prevalence of diabetes did not differ between the hip-fracture groups. In the no-fracture group for the recently untreated patients, the rates of

hip/spine osteoporosis by DXA (28% for women and 17% for men) were similar to those for testing positive by BCT—having either hip osteoporosis or fragile bone strength—(30% for women and 18% for men) and generally were higher than for hip osteoporosis either by DXA or BCT (17% to 19% for women and 12% to 14% for men).

By all metrics and for both sexes, the hip BMD *T*-score (lower of femoral neck and total hip values) for the BCT analysis performed at least as well as DXA testing for assessing risk of hip fracture (Table 2). The hazard ratio (per standard deviation) for the hip BMD *T*-score, which was adjusted for age, race/ethnicity, and BMI, was similar between BCT and DXA and, as expected,[37] was higher than the hazard ratio for the hip/spine BMD *T*-score (by DXA); similar trends were observed among the recently untreated patients. These trends were also reflected in the AUC values for predicting 5-year hip fracture (Table 3). For both sexes, the hazard ratio for patients testing positive for hip osteoporosis trended higher for BCT than DXA, but none of these differences reached statistical significance because of the overlapping confidence intervals (Table 3). The hazard ratios by DXA were all higher for hip osteoporosis than for hip/spine osteoporosis, confirming that combining hip and spine measurements does not increase the predictive ability of the DXA assessment.[37] In general, both sensitivity and specificity for hip osteoporosis were similar between DXA and BCT, again because of the overlapping confidence intervals; sensitivity was generally higher for the women, but specificity was higher for the men; and for DXA, sensitivity was higher for hip/spine osteoporosis, but specificity was higher for hip osteoporosis (Table 3).

Considering only the data derived from the CT scans, femoral strength was associated with hip fracture independent of the hip BMD *T*-score (by BCT) for both sexes. The hazard ratio (per standard deviation) for femoral strength, which was adjusted for age, race/ethnicity, and BMI, was significant for both the women and men and remained so after adjusting also for the (BCT-based) hip BMD *T*-score (Table 2). These trends persisted for the subcohort of recently untreated patients (Table 2). In addition, when the binary classifications for hip osteoporosis (by BCT) and fragile bone strength were included together in the same sex-stratified logistic regression for predicting 5-year hip fracture, the odds ratio for each classification remained significant—hip osteoporosis: 2.4 (95% CI 1.6–3.8) for women and 1.9 (1.1–3.3) for men; fragile bone strength: 2.0 (1.4–3.1) for women and 2.9 (1.7–4.9) for men; interaction terms were not significant.

The main benefit of combining the BCT measurements of the hip BMD *T*-score and femoral strength was to increase sensitivity for predicting fracture, albeit at the expense of specificity (Table 3). The effect appeared slightly larger for the men and depended on how the measurements were combined. Considering only the BCT measurements, sensitivity for the fragile bone strength classification was higher than for the hip osteoporosis classification for both sexes, although for the women, but not the men, that higher sensitivity came with lower specificity, consistent with the larger difference in AUC values found for the men (Table 3). Sensitivity was further increased when these two BCT classifications were combined using an “or” logical—defining patients at “high risk” as those who tested positive for either hip osteoporosis or fragile bone strength. For that classification, 18% more women and 24% more men were correctly identified than when only classifying by (BCT-based) hip osteoporosis (Fig. 2, Table 3); this increase in sensitivity was accompanied by reductions in

specificity of 14% for the women and 7% for the men. By contrast, when the classifications were combined using the more conservative “and” logical—patients at “high risk” were those who tested positive for both hip osteoporosis or fragile bone strength—sensitivity decreased (versu BMD *T*-score alone) for women and men by 5% and 16%, respectively, and specificity increased by 3% and 7%, respectively. Comparing the combined fragile bone strength-or-hip osteoporosis “high-risk” classification by BCT to the hip/spine osteoporosis classification by DXA as a clinical-practice reference, sensitivity for predicting hip fracture was higher for BCT than DXA (women: 0.66 versus 0.59; men: 0.56 versus 0.48), with comparable respective specificity (women: 0.66 versus 0.67; men: 0.76 versus 0.78).

For both sexes, the increase in risk of hip fracture for testing positive by any of the BCT classifications, as reflected by the hazard ratios, was at least as high as for testing positive for osteoporosis by DXA (Table 3). The hazard ratios for the various BCT binary classifications (3.4–3.9 for women; 4.0–4.8 for men, Table 3) consistently trended higher than for the hip osteoporosis classification by DXA (2.9 for women; 3.3 for men), and for the women were significantly higher than for the hip/spine osteoporosis classification by DXA.

Although the main CT imaging-related parameters did not alter these results, there were some noteworthy trends (Table 4). First, there was no evidence of any effects of scan thickness. For example, for the women, hazard ratios for the BCT measurements ranged from 2.2 to 2.9 for thicknesses ≤3.0 mm and from 2.0 to 2.7 for thicknesses >3.0 mm; for the men, hazard ratios were higher for the larger thickness but so too were the values for the DXA measurements, indicating this effect was either a cohort or statistical effect (larger confidence intervals for the smaller sample). Second, there were some trends suggestive of the hazard ratios being higher in non-contrast scans, but the effect did not reach statistical significance. For example, the hazard ratio for strength for women was 3.4 (2.3–4.9) in the subcohort without intravenous contrast versus 2.6 (1.9–3.4) in the subcohort with contrast, whereas the DXA hip BMD remained constant. Consistent with these two results, the two-factor analysis of covariance on the difference in femoral neck BMD between BCT and DXA showed no significant effect on scan thickness ($p = 0.21$) and small but significant effects both for BMI and the presence of intravenous contrast. In particular, increased BMI led to lower BMD *T*-score values for BCT compared with DXA (by 0.012 *T*-score units per unit change in BMI, $p < 0.001$), as did the presence of intravenous contrast (by 0.17 *T*-score units on average, BCT lower than DXA, $p < 0.001$).

Discussion

These results demonstrate that BCT analysis of previously taken routine clinical abdominal or pelvic CT scans is at least as effective as DXA testing for identifying patients at high risk of hip fracture. In all fracture-risk metrics assessed, the hip BMD *T*-score and the femoral strength measurements from BCT each performed as well as the hip BMD *T*-score for DXA testing. Using only the BCT data, combining the classifications of fragile bone strength and hip osteoporosis resulted in higher sensitivity than the clinical-standard hip/spine osteoporosis classification by DXA, at comparable specificity. These findings should apply more generally beyond the FOCUS cohort. All CT and DXA scans in this study were

obtained before the hip fracture and all BCT analyses were performed blinded to the hip fracture status, DXA data, and any clinical factors, resulting in a prospective unbiased test of efficacy. Our large and diverse cohort, drawn from a source population of more than 110,000 women and men and containing almost 1500 patients with incident hip fractures, renders this one of the largest incident hip-fracture outcome studies,[38] supporting external validity. The successful BCT analysis of 86% of the available CT scans, which encompassed 80 different CT scanners and a wide variety of acquisition settings used in 14 hospitals over a 9-year period, supports technical robustness in a broad clinical setting. Importantly for clinical implementation, the BCT analyses were all performed on routine clinical scans, without the need for bone calibration phantoms or any special imaging procedure.

From a basic science perspective, our results help establish that, although correlated with hip BMD (Fig. 2), femoral strength is associated with hip fracture independent of hip BMD for both sexes. Prior BCT studies on incident hip fracture, all conducted in research settings, [20–22,26] have demonstrated consistent trends for a higher gradient of risk for femoral strength than for hip BMD. For example, in the MrOS study of about 3000 men over age 65 years across six regions in the United States, the age-adjusted hazard ratio trended higher for femoral strength (8.0, 2.6–24.3) than for total hip BMD by DXA (4.6, 2.6–8.3);[21] in the AGES study of about 2500 women and 2500 men over age 65 years in Reykjavik, Iceland, the age-adjusted hazard ratio trended higher for femoral strength than for femoral neck BMD (by BCT; DXA was not available) for the women (4.2, 2.6–6.9 versus 2.7, 1.9–3.9) and was similar for the men (3.5, 2.3–5.3 versus 3.7, 2.5–5.6).[20] Both those studies used the VirtuOst software (or its precursor). In a different BCT analysis of the AGES cohort using different software, femoral strength was independently (versus hip BMD) associated with fracture risk for men for simulated stance loading.[22] And studies on prevalent hip fracture have also consistently found that femoral strength is associated with risk of hip fracture.[24,39] In each of these prior studies, the low number of hip fracture cases (≤77 for either sex) has hampered statistical power. Thus, with its almost 10-fold larger number of fracture cases, our new results establish that BCT-derived femoral strength does indeed have an independent association with incident hip fracture relative to hip BMD, for both sexes.

From a clinical perspective, our results establish that BCT can identify patients at high risk of hip fracture on the basis of insufficient femoral strength—those with “fragile bone strength”—and some of these patients do not have BMD-defined osteoporosis at the hip (by BCT). The classification for fragile bone strength was first introduced in 2014,[20] and, as a validated cut-point, is endorsed by the International Society of Clinical Society Densitometry[37] for initiating treatment. We found that the increase in risk of hip fracture for patients with fragile bone strength was comparable to the increase for patients with hip osteoporosis by either BCT or DXA, consistent with our prior BCT studies in a research setting,[20,21] and was significantly higher than for women with hip/spine osteoporosis by DXA (Table 3); femoral strength by BCT is also associated with prevalent fractures of any type.[39] Clinically, because the hip BMD *T*-score is so well established for identifying patients at high risk of hip fracture,[40] we sought to explore how femoral strength can be used *with* the hip BMD *T*-score to improve the BCT assessment. Despite the correlation between hip BMD and femoral strength and the similarity of their AUC values, we found that combining their binary classifications can increase sensitivity—by about 20%—

although at the expense of specificity. Even so, that resulting level of specificity was similar to what we observed for hip/spine osteoporosis by DXA, which is widely used clinically, suggesting this level of specificity is acceptable clinically. Furthermore, we found that the risk of hip fracture for patients testing positive by BCT was comparable to the risk of hip fracture for patients testing positive for hip osteoporosis by DXA. Mechanistically, the increase in sensitivity of femoral strength over the hip BMD *T*-score (for BCT) may reflect the physics of having low femoral strength relative to hip BMD. That effect can occur for such reasons as a deficit in the amount of trabecular bone compared with cortical bone within the femoral neck,[41] local thinning of the cortical bone,[42,43] or a long femoral-neck axis.[44] None of these features are captured in the hip BMD measurement but can be captured in the femoral strength measurement because of its 3D biomechanical nature.

Our results also extend the evidence base of utilizing opportunistic CT for measuring hip BMD from routine clinical CT scans. The concept of ancillary analysis of CT scans for some type of BMD or CT-attenuation measurement is not new.[45–48] With a focus on hip fractures, more recently, we[14–16] and others[9–13] have reported on measuring a DXA-equivalent hip BMD from routine clinical CT scans without the simultaneous use of a calibration phantom. To date, only one of these prior studies has reported on new hip fracture as the clinical outcome.[9] That study demonstrated that hip BMD from CT was indeed associated with new hip fractures, which is consistent with our findings. Our much larger sample size adds a degree of external validity to those findings and helps demonstrate technical robustness.

In analyzing previously taken routine clinical CT scans, our approach to phantomless calibration is noteworthy for two reasons. First, we used each patient's own internal reference tissues as references for calibration, which provides a patient-specific calibration.[17] This approach differs from the "asynchronous"[49] approach, which utilizes a previously measured calibration and therefore provides a machine-specific but not a patient-specific calibration. Our current results demonstrate general robustness of our approach, consistent with two previous studies that validated this approach versus DXA for CT enterography and CT colonography scans.[14,15] The interloperator reanalysis precision for our implementation of BCT is also equivalent to that for traditional quantitative CT (using a calibration phantom).[17] Second, in the current study, we uniquely applied a small adjustment for the presence of any intravenous contrast in the CT scan, based on two independent but consistent observations of such effects in paired sets of CT scans with and without intravenous contrast.[13,32] That adjustment slightly lowers the BMD *T*-score for BCT for scans with intravenous contrast, which contributed to the offset observed between BCT and DXA for the femoral-neck BMD *T*-scores. However, this offset of 0.17 *T*-score units is small compared with the range of precision error (0.82 =±0.41 *T*-score units) for DXA for an individual patient in a clinical setting.[50] Furthermore, this discrepancy more likely reflects a limitation of DXA than BCT. In VirtuOst, the patient's proximal femur is virtually rotated for the BMD measurement into a clinically recommended, standardized orientation that tends to minimize the value of the measured hip BMD. By comparison, it is difficult during DXA scanning in a clinical setting to consistently obtain this ideal femoral orientation—some technicians may be trained in different ways, some may deviate from protocol, or some patients may not be amenable to correct positioning. These issues with

clinical implementation of DXA in a large heterogeneous clinical setting such as FOCUS should typically bias the DXA measurement upwards; they should also double the precision error compared with a research setting,[50] thus diluting DXA's predictive value. Those combined effects well explain the observed offset between the femoral neck measurements as well as the numerically higher values of the hazard ratio and sensitivity for the hip osteoporosis classification for BCT compared with DXA. Thus, although the BMD *T*-scores from BCT did not exactly match those from DXA, we believe the small difference (for the women) primarily reflects the heterogeneous and clinical nature of our DXA data; the larger differences for the men arose primarily from the young-reference values used in Lunar DXA machines not matching those from NHANES-III.

Despite the large size of our hip-fracture cohort, our study has several caveats and limitations. All our patients required an abdominal or pelvic CT exam during the course of their medical care. The prevalence of hip/spine osteoporosis in our recently untreated controls, at 28% in women (average age 73 years) and 17% in men (average age 76 years), is similar to what has been observed for the general Medicare population for women in their 70s (26%) but is threefold higher than for the men in their 70s (5%).[51] This discrepancy for men suggests that men over age 65 years requiring an abdominal or pelvic CT scan may represent an enriched patient group for osteoporosis screening. Related, there were likely various other patient and clinical factors that influenced the association between our main measurements and risk of hip fracture and sensitivity/specificity. Assuming any such factors would likely affect BCT and DXA to a similar extent, our results are best interpreted in terms of the relative performance of BCT with respect to DXA.

There are limitations associated with our sole use of preexisting medical record data in this study. Some patients went on (osteoporosis) treatment as a result of their DXA scan, which confounded comparison with BCT because the CT was within ±3 years of the DXA; this issue required us to use a subcohort to explore, and those results showed some small shifts in hazard ratio values, although those shifts are inconclusive because they may reflect the smaller sample size or true cohort effects. Overall, we lost about 30% of initially selected patients because of either missing or poor-quality data (this number was higher than expected because of a one-time issue with the DXA database). However, this loss of data was unlikely to be a source of bias between BCT and DXA, and despite this loss, the study remained sufficiently powered to demonstrate the significant role of femoral strength after accounting for hip BMD for both sexes. Other limitations relate to data that were not included in the analysis. Although the abdominal CT scans included coverage of the spine, we did not include any spine measurements. The other information omitted was composed of various clinical factors currently used to clinically diagnose osteoporosis,[52,53] such as a prior fragility or vertebral[9] fracture and the FRAX-based absolute risk score.[54] It is not clear how the femoral strength measurement would complement these approaches, another topic of further study.

Related to the use of preexisting clinical data, some limitations derive from the clinical nature of the CT and DXA scans, as noted above. Additionally, because almost all CT scans in our study were acquired at 120 kVp using a standard kernel, our results may not apply to scans acquired at appreciably lower kVp values (eg, 80 kVp) or to scans reconstructed using

sharp or nonstandard kernels—although these settings are not common. Second, we used young-male references for the men and would have obtained different results for the men—primarily lower values of sensitivity—had we used young-female references. However, the main discrepancy in BMD *T*-scores between BCT and DXA for the men likely reflects differences in the young-male reference values between NHANES-III and Lunar DXA machines, highlighting a broader issue of how best to calculate BMD *T*-scores for men.[55] Related, we also noted a significant BMI effect in comparing CT and DXA for hip BMD, consistent with our two prior studies,[14,15] which reflect limitations primarily with DXA.[56] Although resolving these various technical issues is beyond the scope of this study, careful attention to CT-scan quality for individual patients is important for clinical diagnostic applications. Toward that end, BCT is currently performed clinically via a centralized lab model (scans are sent electronically to the lab and a medical report returned within 24 hours) through which quality control can be carefully monitored.

Despite these caveats and limitations, this study has potentially important implications for managing osteoporosis. Our results demonstrate that BCT analysis of previously taken routine abdominal or pelvic CT is at least as effective as DXA testing in a large clinical setting. A recent cost-effectiveness study concluded that BCT analysis, when performed on a dedicated hip CT scan (ordered specifically for BCT analysis), should be cost-effective when the charge for the BCT analysis is set at $100 over the cost of a dedicated hip CT scan.[57] At that pricing, *ancillary* BCT should be even more cost-effective because there is no need for a dedicated hip CT scan. Further, because the ancillary BCT analysis does not require any extra procedure for the patient, a high rate of BCT analysis within the CT population is feasible. Many millions of patients are already scanned with hip CTs each year. For example, in 2007, 24.2 million abdominal CT scans were reported in the US (which includes repeated scans per patient and any age),[18] and in 2016, approximately 10% of individual male and female patients at KPSC aged 65 years or older had at least one hip CT. By contrast, DXA annual screening rates in the Medicare population from 2006–2009 were 9.5% for women and 1.7% for men[8] and are unlikely to have risen since then.[3,5,6] Thus, if BCT analysis were routinely used in women and men over age 65 years who have an available hip CT when a recommended DXA test is either missing, not available, or inconvenient, or perhaps even if used on dedicated hip CT scans,[57] such a “second front” in osteoporosis screening, could appreciably increase the number of patients identified as being at high risk of hip fracture. This additional screening in turn could appreciably reduce the number of hip fractures.[4]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by unrestricted investigator-initiated research funds from Amgen and Merck and by NIH grant R44 AR057616 (PI: DLK). The funding agencies had no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript. TMK, DLK, and DCL have financial interests in O.N. Diagnostics and both they and the company may benefit from the results of this work.

Authors' roles: Study design: TMK, DLK, HF, and ALA. Study conduct: TMK, DLK, DCL, HF, and ALA. Data collection: DCL and HF. Data analysis: HF. Data interpretation: HF, DLK, DMB, MLB, SF, SK, ESO, ESS, TMK, and ALA. Drafting manuscript: TMK, ALA, and HF. Revising manuscript content: DLK, DCL, DMB, MLB, SF, SK, ESO, and ESS. Approving final version of manuscript: ALA, HF, DLK, DCL, DMB, MLB, SF, SK, ESO, ESS, and TMK. HF takes responsibility for the integrity of the data analysis.

References

1. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. J Bone Miner Res 2007;22(3): 465–75. [PubMed: 17144789]
2. Bass E, French DD, Bradham DD. A national perspective of Medicare expenditures for elderly veterans with hip fractures. J Am Med Dir Assoc 2008;9(2):114–9. [PubMed: 18261704]
3. Lewiecki EM, Laster AJ, Miller PD, Bilezikian JP. More bone density testing is needed, not less. J Bone Miner Res 2012;27(4):739–42. [PubMed: 22383457]
4. Siris ES, Pasquale MK, Wang Y, Watts NB. Estimating bisphosphonate use and fracture reduction among US women aged 45 years and older, 2001–2008. J Bone Miner Res 2011;26(1):3–11. [PubMed: 20662073]
5. Golob AL, Laya MB. Osteoporosis: screening, prevention, and management. Med Clin N Am 2015;99(3):587–606. [PubMed: 25841602]
6. King AB, Fiorentino DM. Medicare payment cuts for osteoporosis testing reduced use despite tests' benefit in reducing fractures. Health Aff (Millwood) 2011;30(12):2362–70. [PubMed: 22147865]
7. Lim SY, Lim JH, Nguyen D, et al. Screening for osteoporosis in men aged 70 years and older in a primary care setting in the United States. Am J Mens Health. 2013;7(4):350–4. [PubMed: 23440083]
8. Zhang J, Delzell E, Zhao H, et al. Central DXA utilization shifts from office-based to hospital-based settings among Medicare beneficiaries in the wake of reimbursement changes. J Bone Miner Res 2012;27(4):858–64. [PubMed: 22190195]
9. Lee SJ, Anderson PA, Pickhardt PJ. Predicting future hip fractures on routine abdominal CT using opportunistic osteoporosis screening measures: a matched case-control study. AJR Am J Roentgenol 2017;209(2):395–402. [PubMed: 28570093]
10. Ziemlewicz TJ, Maciejewski A, Binkley N, Brett AD, Brown JK, Pickhardt PJ. 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 2016;31(10):1835–40. [PubMed: 27082831]
11. Ziemlewicz TJ, Binkley N, Pickhardt PJ. Opportunistic osteoporosis screening: addition of quantitative CT bone mineral density evaluation to CT colonography. J Am Coll Radiol 2015;12(10): 1036–41. [PubMed: 26435117]
12. Pickhardt PJ, Bodeen G, Brett A, Brown JK, Binkley N. Comparison of femoral neck BMD evaluation obtained using Lunar DXA and QCT with asynchronous calibration from CT colonography. J Clin Densitom 2015;18(1):5–12. [PubMed: 24880495]
13. Ziemlewicz TJ, Maciejewski A, Binkley N, Brett AD, Brown JK, Pickhardt PJ. Direct comparison of unenhanced and contrast-enhanced CT for opportunistic proximal femur bone mineral density measurement: implications for osteoporosis screening. AJR Am J Roentgenol 2016;206(4):694–8. [PubMed: 26866336]
14. Weber NK, Fidler JL, Keaveny TM, et al. Validation of a CT-derived method for osteoporosis screening in IBD patients undergoing contrast-enhanced CT enterography. Am J Gastroenterol 2014;109(3):401–8. [PubMed: 24445572]
15. Fidler JL, Murthy NS, Khosla S, et al. Comprehensive assessment of osteoporosis and bone fragility with CT colonography. Radiology 2016;278(1):172–80. [PubMed: 26200602]
16. Schwaiger BJ, Kopperdahl DL, Nardo L, et al. Vertebral and femoral bone mineral density and bone strength in prostate cancer patients assessed in phantomless PET/CT examinations. Bone. 2017;101:62–9. [PubMed: 28442297]

17. Lee DC, Hoffmann PF, Kopperdahl DL, Keaveny TM. Phantomless calibration of CT scans for measurement of BMD and bone strength —inter-operator reanalysis precision. Bone. 2017;103:325–33. [PubMed: 28778598]
18. Berrington de Gonzalez A, Mahesh M, Kim KP, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. Arch Intern Med 2009;169(22):2071–7. [PubMed: 20008689]
19. Keaveny TM. Biomechanical computed tomography-noninvasive bone strength analysis using clinical computed tomography scans. Ann N Y Acad Sci 2010;1192:57–65. [PubMed: 20392218]
20. Kopperdahl DL, Aspelund T, Hoffmann PF, et al. Assessment of incident spine and hip fractures in women and men using finite element analysis of CT scans. J Bone Miner Res 2014;29(3):570–80. [PubMed: 23956027]
21. Orwoll ES, Marshall LM, Nielson CM, et al. Finite element analysis of the proximal femur and hip fracture risk in older men. J Bone Miner Res 2009;24(3):475–83. [PubMed: 19049327]
22. Keyak JH, Sigurdsson S, Karlsdottir G, et al. Male-female differences in the association between incident hip fracture and proximal femoral strength: a finite element analysis study. Bone. 2011;48(6):1239–45. [PubMed: 21419886]
23. Bessho M, Ohnishi I, Matsumoto T, et al. 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. 2009;45(2):226–31. [PubMed: 19398043]
24. Nishiyama KK, Ito M, Harada A, Boyd SK. Classification of women with and without hip fracture based on quantitative computed tomography and finite element analysis. Osteoporos Int 2014;25(2):619–26. [PubMed: 23948875]
25. Qasim M, Farinella G, Zhang J, et al. Patient-specific finite element estimated femur strength as a predictor of the risk of hip fracture: the effect of methodological determinants. Osteoporos Int 2016;27(9):2815–22. [PubMed: 27108118]
26. Falcinelli C, Schileo E, Balistreri L, et al. 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. 2014;67:71–80. [PubMed: 25014885]
27. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Philadelphia: Lippincott Williams & Wilkins; 2008.
28. Keaveny TM, Kopperdahl DL, Melton LJ, 3rd, et al. Age-dependence of femoral strength in white women and men. J Bone Miner Res 2010;25(5):994–1001. [PubMed: 19874201]
29. Johannesdottir F, Thrall E, Muller J, Keaveny TM, Kopperdahl DL, Bouxsein ML. Comparison of non-invasive assessments of strength of the proximal femur. Bone. 2017;105:93–102. [PubMed: 28739416]
30. Schousboe JT, Shepherd JA, Bilezikian JP, Baim S. Executive summary of the 2013 International Society for Clinical Densitometry Position Development Conference on bone densitometry. J Clin Densitom 2013;16(4):455–66. [PubMed: 24183638]
31. Looker AC, Wahner HW, Dunn WL, et al. Updated data on proximal femur bone mineral levels of US adults. Osteoporos Int 1998;8(5):468–89. [PubMed: 9850356]
32. Bauer JS, Henning TD, Mueller D, Lu Y, Majumdar S, Link TM. Volumetric quantitative CT of the spine and hip derived from contrast-enhanced MDCT: conversion factors. Am J Roentgenol 2007;188(5):1294–301. [PubMed: 17449773]
33. Dell RM, Greene D, Anderson D, Williams K. Osteoporosis disease management: what every orthopaedic surgeon should know. J Bone Joint Surg Am 2009;91 Suppl 6:79–86. [PubMed: 19884415]
34. Qaseem A, Forciea MA, McLean RM, Denberg TD. Treatment of low bone density or osteoporosis to prevent fractures in men and women: a clinical practice guideline update from the American College of Physicians. Ann Intern Med 2017;166(11):818–39. [PubMed: 28492856]
35. Breslow NE, Lumley T, Ballantyne CM, Chambless LE, Kulich M. Using the whole cohort in the analysis of case-cohort data. Am J Epidemiol 2009;169(11):1398–405. [PubMed: 19357328]
36. Thiebaut AC, Benichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. Stat Med 2004;23(24):3803–20. [PubMed: 15580597]

37. Zysset P, Qin L, Lang T, et al. 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 2015;18(3):359–92. [PubMed: 26277852]
38. Johansson H, Kanis JA, Oden A, et al. A meta-analysis of the association of fracture risk and body mass index in women. J Bone Miner Res 2014;29(1):223–33. [PubMed: 23775829]
39. Amin S, Kopperdahl DL, Melton LJ, 3rd, et al. Association of hip strength estimates by finite-element analysis with fractures in women and men. J Bone Miner Res 2011;26(7):1593–600. [PubMed: 21305605]
40. Kanis JA, Johnell O, Oden A, et al. The use of multiple sites for the diagnosis of osteoporosis. Osteoporos Int 2006;17(4):527–34. [PubMed: 16402164]
41. Nawathe S, Akhlaghpour H, Bouxsein ML, Keaveny TM. Microstructural failure mechanisms in the human proximal femur for sideways fall loading. J Bone Miner Res 2014;29(2):507–15. [PubMed: 23832419]
42. Poole KE, Skingle L, Gee AH, et al. Focal osteoporosis defects play a key role in hip fracture. Bone. 2017;94:124–34. [PubMed: 27777119]
43. Museyko O, Bousson V, Adams J, Laredo J, Engelke K. QCT of the proximal femur—which parameters should be measured to discriminate hip fracture? Osteoporos Int 2016;27(3):1137–47. [PubMed: 26415934]
44. Leslie WD, Lix LM, Morin SN, et al. Hip axis length is a FRAX- and bone density-independent risk factor for hip fracture in women. J Clin Endocrinol Metab 2015;100(5):2063–70. [PubMed: 25751108]
45. Pickhardt PJ, Lee LJ, del Rio AM, et al. Simultaneous screening for osteoporosis at CT colonography: bone mineral density assessment using MDCT attenuation techniques compared with the DXA reference standard. J Bone Miner Res 2011;26(9):2194–203. [PubMed: 21590738]
46. Summers RM, Baecher N, Yao J, et al. Feasibility of simultaneous computed tomographic colonography and fully automated bone mineral densitometry in a single examination. J Comput Assist Tomogr 2011;35(2):212–6. [PubMed: 21412092]
47. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N. Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. Ann Intern Med 2013;158(8):588–95. [PubMed: 23588747]
48. Mueller DK, Kutscherenko A, Bartel H, Vlassenbroek A, Ourednicek P, Erckenbrecht J. Phantom-less QCT BMD system as screening tool for osteoporosis without additional radiation. Eur J Radiol 2011;79(3):375–81. [PubMed: 20223609]
49. Brown JK, Timm W, Bodeen G, et al. Asynchronously calibrated quantitative bone densitometry. J Clin Densitom 2017;20(2): 216–25. [PubMed: 26781430]
50. Kiebzak GM, Faulkner KG, Wacker W, Hamdy R, Seier E, Watts NB. Effect of precision error on T-scores and the diagnostic classification of bone status. J Clin Densitom 2007;10(3):239–43. [PubMed: 17451984]
51. Wright NC, Looker AC, Saag KG, et al. 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 2014;29(11):2520–6. [PubMed: 24771492]
52. Camacho PM, Petak SM, Binkley N, et al. American Association of Clinical Endocrinologists and American College of Endocrinology clinical practice guidelines for the diagnosis and treatment of postmenopausal osteoporosis - 2016. Endocr Pract 2016;22 Suppl 4:1–42.
53. Cosman F, de Beur SJ, LeBoff MS, et al. Clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 2014;25(10): 2359–81. [PubMed: 25182228]
54. Kanis JA, Oden A, Johnell O, et al. 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 2007;18(8):1033–46. [PubMed: 17323110]
55. Srinivasan B, Kopperdahl DL, Amin S, et al. Relationship of femoral neck areal bone mineral density to volumetric bone mineral density, bone size, and femoral strength in men and women. Osteoporos Int 2012;23(1):155–62. [PubMed: 22057550]

56. Bolotin HH. DXA in vivo BMD methodology: an erroneous and misleading research and clinical gauge of bone mineral status, bone fragility, and bone remodelling. Bone. 2007;41(1):138–54. [PubMed: 17481978]

57. Agten CA, Ramme AJ, Kang S, Honig S, Chang G. Cost-effectiveness of virtual bone strength testing in osteoporosis screening programs for postmenopausal women in the United States. Radiology 2017;285(2):506–17. [PubMed: 28613988]

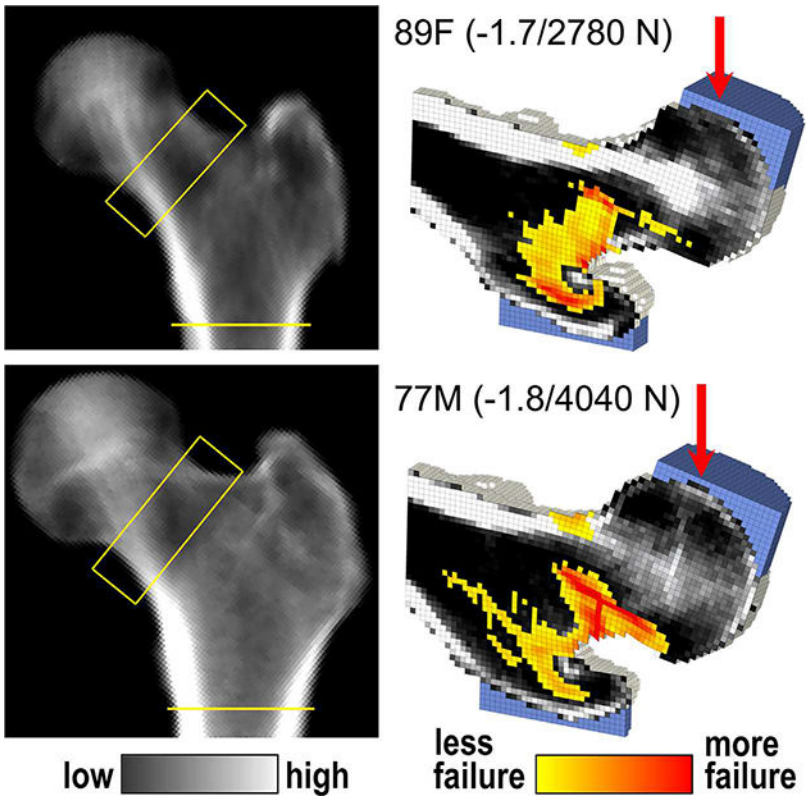

Fig. 1.

Example of the DXA-equivalent hip BMD measurements and the finite element analysis–derived virtual stress testing for femoral strength for two patients. The gray scale shows the level of BMD and the yellow-red colors represent regions of failed tissue in the virtual stress testing; the labels denote age and sex (hip BMD *T*-score/femoral strength). Both patients tested negative for hip osteoporosis (BMD *T*-score >–2.5), but the top patient tested positive for fragile bone strength (femoral strength <3000 N for women, <3500 N for men).

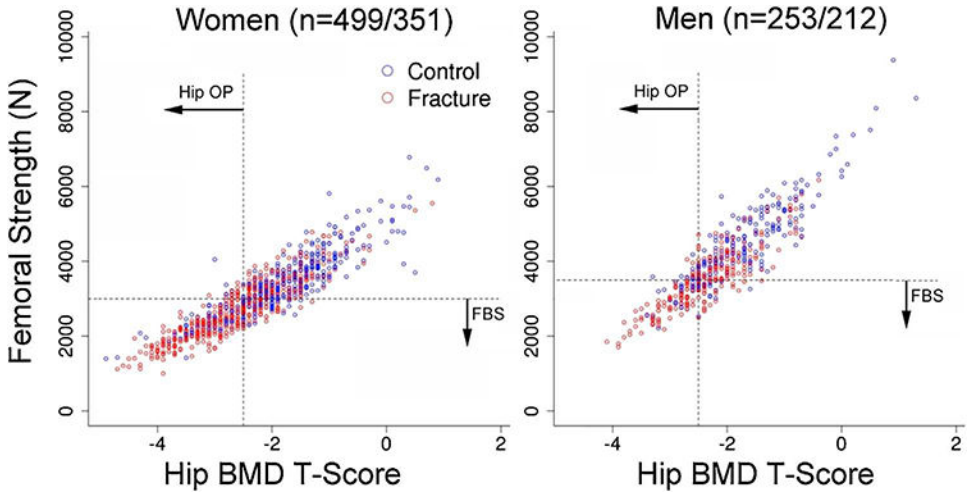

Fig. 2.

Relation between femoral strength and the hip BMD *T*-score, both measured by BCT, for the recently untreated patients with full 5-year fracture-outcome data. Dashed lines indicate the cut-points used to define hip osteoporosis (Hip OP) and fragile bone strength (FBS). Numbers in parentheses show the numbers of fractures/no-fracture controls in the plots. See Table 3 for the sensitivity and specificity data corresponding to these plots.

Table 1.

Patient Characteristics for All Patients and Those Not Treated Recently With Any Osteoporosis Medication, With and Without Hip Fracture (Mean [SD] or Number [%])[a]

	Women				Men			
	All patients		Recently untreated		All patients		Recently untreated	
	No-fracture	Fracture	No-fracture	Fracture	No-fracture	Fracture	No-fracture	Fracture
Number	1019	903	769	595	458	403	404	308
Age (years)	74.0 (7.21)	79.5 (7.37)	73.4 (6.98)	79.1 (7.59)	75.7 (6.75)	80.3 (6.66)	75.5 (6.76)	80.1 (6.63)
Body mass index (kg/m^2)	27.9 (6.22)	25.0 (5.22)	28.7 (6.36)	25.5 (5.32)	27.7 (4.87)	25.6 (4.19)	27.9 (4.97)	25.8 (4.36)
Diabetes	293 (29%)	253 (28%)	233 (30%)	174 (29%)	145 (32%)	142 (35%)	129 (32%)	111 (36%)
Race/ethnicity								
Asian	62 (6%)	38 (4%)	41 (5%)	19 (3%)	51 (11%)	12 (3%)	41 (10%)	10 (3%)
Black	126 (12%)	53 (6%)	110 (14%)	39 (7%)	44 (10%)	26 (7%)	43 (11%)	20 (7%)
Hispanic	288 (28%)	171 (19%)	212 (28%)	110 (19%)	106 (23%)	75 (19%)	96 (24%)	59 (19%)
White/other	543 (53%)	641 (71%)	406 (53%)	427 (72%)	257 (56%)	290 (72%)	224 (55%)	219 (71%)
DXA measurements								
Hip BMD *T*-score	−1.8 (0.98)	−2.6 (0.90)	−1.6 (0.98)	−2.4 (0.94)	−1.5 (1.13)	−2.5 (0.87)	−1.3 (1.10)	−2.3 (0.89)
Hip/spine BMD *T*-score	−2.1 (1.09)	−2.8 (1.05)	−1.9 (1.08)	−2.7 (1.09)	−1.6 (1.15)	−2.6 (0.91)	−1.5 (1.11)	−2.4 (0.91)
Hip osteoporosis	239 (24%)	508 (56%)	134 (17%)	293 (49%)	91 (20%)	205 (51%)	55 (14%)	126 (41%)
Hip/spine osteoporosis	374 (37%)	572 (63%)	213 (28%)	336 (57%)	109 (24%)	222 (55%)	68 (17%)	142 (46%)
BCT measurements								
Hip BMD *T*-score	−1.9 (0.92)	−2.6 (0.84)	−1.7 (0.93)	−2.5 (0.88)	−1.7 (0.83)	−2.4 (0.66)	−1.6 (0.82)	−2.3 (0.66)
Femoral strength (N)	3400 (840)	2730 (750)	3510 (860)	2810 (780)	4470 (1100)	3490 (810)	4600 (1090)	3620 (820)
Hip osteoporosis	252 (25%)	541 (60%)	148 (19%)	317 (53%)	74 (16%)	212 (53%)	49 (12%)	135 (44%)
Fragile bone strength	331 (33%)	594 (66%)	211 (27%)	360 (61%)	84 (18%)	214 (53%)	56 (14%)	142 (46%)
High risk (hip OP or FBS)	364 (36%)	627 (69%)	228 (30%)	379 (64%)	102 (22%)	252 (63%)	71 (18%)	167 (54%)

DXA = dual-energy X-ray absorptiometry; BMD = bone mineral density; BCT = biomechanical computed tomography; OP = osteoporosis; FBS = fragile bone strength.

[a]Those without hip fracture were randomly sampled from the study population, and any hip-fracture cases (~2%) removed.

Age and body mass index were measured at time of the CT exam. Diabetes rates did not differ between any case and control groups ($p = 0.25$–0.72). All other measurements were significantly different between case and control groups, at the $p < 0.0001$ level (Kruskal-Wallis for continuous variables and chi-square for proportions). Hip BMD *T*-score is the lowest available *T*-score for the femoral-neck or total-hip regions for DXA; hip/spine BMD *T*-score is the lower of the hip or total-spine BMD *T*-scores. For BCT, “high risk” denotes the presence of either hip osteoporosis or fragile bone strength.

Table 2.

Hazard Ratios (95% Confidence Intervals [CI]) for Hip Fracture for the BCT and DXA Measurements for All Patients and for Patients Not Recently Treated With Any Osteoporosis Medication

	Women			Men		
	N	Hazard ratio (95% CI)		*N*	Hazard ratio (95% CI)	
All patients						
BCT						
Hip BMD *T*-score	1922	2.1	(1.8–2.5)[a]	861	2.8	(2.1 −3.8)
Femoral strength	1922	2.8	(2.2–3.5)[a,b]	861	2.8	(2.1 −3.7)
Femoral strength (with BMD)	1922	2.1	(1.4–3.2)	861	2.7	(1.6–4.6)
DXA						
Hip BMD *T*-score	1922	2.1	(1.7–2.5)	861	2.5	(2.0–3.2)
Hip/spine BMD T-score	1922	1.6	(1.4–1.9)	861	2.1	(1.6–2.7)
Recently untreated patients						
BCT						
Hip BMD *T*-score	1364	2.4	(2.0–2.9)[a]	712	2.7	(2.0–3.8)
Femoral strength	1364	3.6	(2.8–4.6)[a,b]	712	3.0	(2.2–4.3)
Femoral strength (with BMD)	1364	2.4	(1.5–4.0)	712	3.3	(1.8–5.9)
DXA						
Hip BMD *T*-score	1364	2.3	(1.9–2.9)	712	2.4	(1.9–3.1)
Hip/spine BMD *T*-score	1364	1.7	(1.4–2.0)	712	2.0	(1.6–2.6)

BCT = biomechanical computed tomography; DXA = dual-energy X-ray absorptiometry; BMD = bone mineral density.

All hazard ratios denote the increase in risk of hip fracture per unit decrease of one sex-pooled standard deviation of the measurement and are adjusted for age, race/ethnicity, and body mass index. Femoral strength (with BMD) was additionally adjusted for the hip BMD *T*-score (by BCT). *N* is the combined number of cases and controls (approximately 1:1 ratio). See Table 1 for additional legends.

[a] Significantly different than for the DXA hip/spine BMD *T*-score ($p < 0.05$, at least). Comparisons only made between BCT and DXA.

[b] Significantly different than for the DXA hip BMD *T*-score ($p < 0.05$, at least). Comparisons only made between BCT and DXA.

Table 3.

Sensitivity, Specificity, Area-Under-the-Curve (AUC), and Hazard Ratios, With 95% Confidence Intervals (CI), by BCT and DXA for Different Binary Classifications for Identifying Who Will Have a New Hip Fracture Within 5 Years of the Respective Scan (Recently Untreated Patients Only)

Classification	*N*		Sensitivity (95% CI)		Specificity (95% CI)			AUC	*N*	Hazard ratio (95% CI)	
Women											
BCT											
Hip osteoporosis	850	0.56	[277/499]	(0.51 −0.60)	0.77	[270/351]	(0.72–0.81)	0.72	1364	3.7	(2.7–5.1)[a]
Fragile bone strength	850	0.63	[316/499]	(0.59–0.68)	0.69	[242/351]	(0.64–0.74)	0.73	1364	3.4	(2.5–4.6)[a]
High risk (OP or FBS)	850	0.66	[330/499]	(0.62–0.70)	0.66	[233/351]	(0.61 −0.71)	0.73[b]	1364	3.4	(2.5–4.6)[a]
High risk (OP and FBS)	850	0.53	[263/499]	(0.48–0.57)	0.79	[279/351]	(0.75–0.84)	0.73[b]	1364	3.9	(2.8–5.4)[a]
DXA											
Hip osteoporosis	882	0.52	[245/475]	(0.47–0.56)	0.77	[313/407]	(0.73–0.81)	0.72	1364	2.9	(2.1 −4.0)
Hip/spine osteoporosis	882	0.59	[281/475]	(0.55–0.64)	0.67	[271/407]	(0.62–0.71)	0.70	1364	2.2	(1.6–3.0)
Men											
BCT											
Hip osteoporosis	465	0.45	[115/253]	(0.39–0.52)	0.82	[173/212]	(0.76–0.87)	0.71	712	4.0	(2.4–6.6)
Fragile bone strength	465	0.48	[122/253]	(0.42–0.55)	0.82	[174/212]	(0.76–0.87)	0.75	712	4.1	(2.5–6.8)
High risk (OP or FBS)	465	0.56	[142/253]	(0.50–0.62)	0.76	[161/212]	(0.70–0.82)	0.75[b]	712	4.0	(2.6–6.3)
High risk (OP and FBS)	850	0.38	[95/253]	(0.32–0.44)	0.88	[186/212]	(0.83–0.92)	0.75[b]	712	4.8	(2.6–9.0)
DXA											
Hip osteoporosis	475	0.43	[109/252]	(0.37–0.50)	0.83	[185/223]	(0.77–0.88)	0.73	712	3.3	(2.0–5.3)
Hip/spine osteoporosis	475	0.48	[122/252]	(0.42–0.55)	0.78	[173/223]	(0.72–0.83)	0.72	712	3.3	(2.1 −5.1)

BCT = biomechanical computed tomography; OP = osteoporosis; FBS = fragile bone strength; DXA = dual-energy X-ray absorptiometry.

Hazard ratios denote the increase in risk of hip fracture for testing positive for the binary classification and are adjusted for age, race/ethnicity, and body mass index. All no-fracture patients for sensitivity/specificity analysis require a 5-year follow-up from their scan and thus the sample size differs between BCT and DXA and is lower than for the hazard ratio analysis, which can adjust for follow-up time. See Tables 1 and 2 for additional legends.

[a] Significantly different than for the DXA hip/spine BMD *T*-score ($p < 0.05$, at least). Comparisons only made between BCT and DXA.

[b] AUC values for the high-risk classifications include hip BMD and femoral strength in the same model and therefore are the same for the “and” and “or” logicals.

Table 4.

Effects of Factors Related to CT Scan Acquisition Settings and Hazard Ratios (95% Confidence Intervals [CI]) for Hip Fracture for the BCT and DXA Measurements for Various Subcohorts Stratified by CT Scan Parameters

	Women		Men	
	N	**Hazard ratio (95% CI)**	***N***	**Hazard ratio (95% CI)**
Contrast-enhanced CT scan				
BCT				
Hip BMD *T*-score	1293	2.0 (1.6–2.4)	471	2.7 (1.8–3.9)
Femoral strength	1293	2.6 (1.9–3.4)[a]	471	2.8 (1.9–4.1)
Femoral strength (with BMD)	1293	1.9 (1.2–3.1)	471	3.4 (1.6–7.2)
DXA				
Hip BMD *T*-score	1293	2.1 (1.7–2.6)	471	2.3 (1.8–3.1)
Hip/spine BMD T-score	1293	1.8 (1.5–2.1)	471	2.3 (1.6–3.3)
Non-contrast-enhanced CT scan				
BCT				
Hip BMD *T*-score	629	2.5 (1.9–3.2)[a]	390	4.0 (2.3–6.9)[a]
Femoral strength	629	3.4 (2.3–4.9)[a,b]	390	3.4 (2.1 −5.5)[a]
Femoral strength (with BMD)	629	2.4 (1.1 −5.1)	390	2.3 (1.1 −5.1)
DXA				
Hip BMD *T*-score	629	2.0 (1.5–2.7)	390	3.6 (2.3–5.7)
Hip/spine BMD *T*-score	629	1.5 (1.2–1.9)	390	1.8 (1.2–2.7)
CT scan thickness <3.0 mm				
BCT				
Hip BMD *T*-score	1353	2.2 (1.7–2.7)	604	2.6 (1.9–3.7)
Femoral strength	1353	2.9 (2.2–3.9)[a]	604	2.6 (1.9–3.6)
Femoral strength (with BMD)	1353	2.0 (1.2–3.4)	604	2.7 (1.5–4.9)
DXA				
Hip BMD *T*-score	1353	2.1 (1.7–2.7)	604	2.5 (1.8–3.4)
Hip/spine BMD *T*-score	1353	1.7 (1.4–2.1)	604	2.2 (1.6–3.2)
CT scan thickness >3.0 mm				
BCT				
Hip BMD *T*-score	569	2.0 (1.6–2.7)	257	4.7 (2.4–9.4)
Femoral strength	569	2.7 (1.8–4.0)[a]	257	5.7 (3.0–10.7)[a]
Femoral strength (with BMD)	569	2.5 (1.2–5.2)	257	5.2 (1.7–16.0)
DXA				
Hip BMD *T*-score	569	1.9 (1.4–2.5)	257	3.5 (2.5–4.7)
Hip/spine BMD *T*-score	569	1.5 (1.1–1.9)	257	2.5 (1.7–3.7)

BCT = biomechanical computed tomography; BMD = bone mineral density; DXA = dual-energy X-ray absorptiometry. See Tables 1 and 2 for legends.

[a]Significantly different than for the DXA hip/spine BMD T-score ($p < 0.05$, at least). Comparisons only made between BCT and DXA.

[b] Significantly different than for the DXA hip BMD T-score ($p < 0.05$, at least). Comparisons only made between BCT and DXA.