



Published in final edited form as:

*J Bone Miner Res.* 2014 March ; 29(3): 570–580. doi:10.1002/jbmr.2069.

## Assessment of Incident Spine and Hip Fractures in Women and Men using Finite Element Analysis of CT Scans

David L. Kopperdahl, PhD<sup>1</sup>, Thor Aspelund, PhD<sup>2,4</sup>, Paul F. Hoffmann, BS<sup>1</sup>, Sigurdur Sigurdsson, MSc<sup>2</sup>, Kristin Siggeirsdottir, MS<sup>2</sup>, Tamara B. Harris, MD, MS<sup>3</sup>, Vilundur Gudnason, MD, PhD<sup>2,4</sup>, and Tony M. Keaveny, PhD<sup>1,5</sup>

<sup>1</sup>O.N. Diagnostics, Berkeley, CA, USA <sup>2</sup>Icelandic Heart Association Research Institute, Kopavogur, Iceland <sup>3</sup>Intramural Research Program, National Institute on Aging, Bethesda, MD, USA <sup>4</sup>University of Iceland, Reykjavik, Iceland <sup>5</sup>Departments of Mechanical Engineering and Bioengineering, University of California, Berkeley, CA, USA

### Abstract

Finite element analysis of computed tomography (CT) scans provides non-invasive estimates of bone strength at the spine and hip. To further validate such estimates clinically, we performed a five-year case-control study of 1110 women and men over age 65 from the AGES-Reykjavik cohort (case = incident spine or hip fracture; control = no incident spine or hip fracture, respectively). From the baseline CT scans, we measured femoral and vertebral strength, as well as bone mineral density (BMD) at the hip (areal BMD only) and lumbar spine (trabecular volumetric BMD only). We found that, for incident radiographically-confirmed spine fractures (n=167), the age-adjusted odds ratio for vertebral strength was significant for women (2.8, 95% CI: 1.8–4.3) and men (2.2, 95% CI: 1.5–3.2), and for men, remained significant (p=0.01) independent of vertebral trabecular volumetric BMD. For incident hip fractures (n=171), the age-adjusted odds ratio for femoral strength was significant for women (4.2, 95% CI: 2.6–6.9) and men (3.5, 95% CI: 2.3–5.3) and remained significant after adjusting for femoral neck areal BMD in women and for total hip areal BMD in both sexes; fracture classification improved for women by combining femoral strength with femoral neck areal BMD (p=0.002). For both sexes, the probabilities of spine and hip fractures were similarly high at the BMD-based interventional thresholds for osteoporosis and at corresponding pre-established thresholds for “fragile bone strength” (spine: women  $\leq 4,500$  N, men  $\leq 6,500$  N; hip: women  $\leq 3,000$  N, men  $\leq 3,500$  N). Since it is well established that individuals over age 65 who have osteoporosis at the hip or spine by BMD criteria should be considered at high risk of fracture, these results indicate that individuals who have “fragile bone strength” at the hip or spine should also be considered at high risk of fracture.

### Keywords

osteoporosis; fracture risk assessment; biomechanics; bone QCT

---

Address all correspondence to: David Kopperdahl, 2150 Shattuck Ave., Suite 610, Berkeley, CA 94703; phone 510-204-0688; fax 510-356-4349; david.kopperdahl@ondiagnosics.com.

**Disclosures:** DLK and PFH are employees of O.N. Diagnostics. DLK, PFH and TMK each have a financial interest in O.N. Diagnostics, and they and the company may benefit from the results of this research. All other authors state that they have no conflicts of interest.

Author contributions included study conception and design (DLK, TA, TH, VG, TMK), acquisition of data (PFH, SS, KS), statistical analysis (TA), interpretation of data (DLK, TA, TH, VG, TMK), drafting manuscript (DLK, TMK), and manuscript revision and approval (DLK, TA, PFH, SS, KS, TH, VG, TMK). DLK takes responsibility for the integrity of the data analysis.

## INTRODUCTION

The increasing size of the elderly population, coupled with insufficient rates of screening,<sup>(1, 2)</sup> has renewed calls to increase the number of people who are tested for osteoporosis and assessed for fracture risk.<sup>(3)</sup> While DXA is the clinical standard for such testing, finite element analysis (FEA) of computed tomography (CT) scans can now be used clinically to assess fracture risk. FEA-derived strength estimates have been validated in cadaver studies by numerous groups for both the spine<sup>(4-6)</sup> and hip.<sup>(7-12)</sup> These strength estimates have been clinically validated for prediction of incident clinical spine<sup>(13)</sup> and hip<sup>(14, 15)</sup> fractures in men, and incident hip fractures in women.<sup>(15)</sup> Associations have also been shown in women between FEA and both prevalent spine fractures<sup>(5, 16-19)</sup> and any prevalent osteoporotic fracture.<sup>(20)</sup> However, validation studies have not yet been reported for prediction of incident spine fractures in women, and no single FEA methodology has yet been validated in a single study for both incident spine and hip fractures in both sexes. Further, while clinical guidelines for interpreting bone mineral density (BMD) are well established ( $T \leq -2.5$  for defining osteoporosis, for example), such guidelines remain to be validated for FEA-derived measures of strength.

Addressing these limitations, we sought to further validate FEA for predicting incident spine and hip fractures in both women and men, and with clinical translation in mind, to prospectively evaluate pre-established FEA strength-based interventional thresholds for identifying those at high risk of fracture. Since the same CT scans used for FEA can also be used to measure clinically established BMD measures at the spine (volumetric BMD of the trabecular bone)<sup>(21)</sup> and hip (DXA-equivalent total hip and femoral neck areal BMD),<sup>(15, 22)</sup> we also assessed these BMD measures to facilitate clinical interpretation of the FEA results.

## METHODS

For our study design, we used a case-control approach with separate spine and hip arms (Figure 1). For the spine and hip arms, the cases had an incident spine or hip fracture, respectively, during a five-year observation period, whereas the controls did not have an incident spine or hip fracture, respectively. The study participants were drawn from community dwelling women and men in Reykjavik, Iceland who were enrolled in the ongoing Age, Gene/Environment Susceptibility Reykjavik observational study (AGES-Reykjavik).<sup>(23)</sup> All participants in that larger study were scheduled to have a CT exam at a baseline time point, and again after five years, all scans taken on the same CT machine. We evaluated the finite element-estimated bone strength (in Newtons) and load-to-strength ratio measured from the baseline CT scans of the spine and hip for prediction of spine and hip fractures, respectively. We also investigated BMD measures that are currently used clinically, specifically, volumetric vertebral trabecular BMD (in  $\text{mg}/\text{cm}^3$ ) in the spine arm, and femoral neck and total hip areal BMD (in  $\text{g}/\text{cm}^2$ ) in the hip arm. Since DXA was not available, we used the baseline CT scans to measure a DXA-equivalent hip areal BMD (see Appendix A). All CT analyses, including finite element analyses, were performed blinded to fracture status, and the interventional strength thresholds were developed on an independent cohort of 1459 women and men before this analysis.

### Analysis Sample

For the spine arm, incident vertebral fracture cases included both clinical (identified from patient medical records) and morphologic (identified from the study CT scans) fractures. Based on a search of medical records for appropriate ICD-10 codes,<sup>(24)</sup> we identified 69 AGES-Reykjavik participants who had a potential clinical incident spine fracture, 30 of whom were subsequently excluded because the fracture could not be confirmed radiographically via CT scans. Based on reading the baseline and five-year follow-up CT

scans for a random selection of 897 AGES-Reykjavik participants who had both such CT exams, we identified an additional 129 participants who had a morphologic incident spine fracture between baseline and follow-up. One of these cases was subsequently excluded because the baseline CT scan was unsuitable for FEA, resulting in 167 spine fracture cases (39 clinical plus 128 morphologic) for statistical analysis. From that same random selection, we identified 676 controls, namely, participants who displayed a radiographic absence of incident vertebral fracture between baseline and five-year follow-up CT (hip fracture status was not considered in the spine arm; and vice versa). The remaining 93 subjects were excluded from further analysis in the spine arm due to lack of a readable CT scout view (at baseline or follow-up) needed to adjudicate vertebral fracture. The final spine arm comprised 497 women and 346 men, all Caucasian, the age at baseline spanning 66–93 years.

For the hip arm, again based on a medical record search for appropriate ICD-10 codes,<sup>(24)</sup> we identified 180 participants in the AGES-Reykjavik study who had a clinical diagnosis of a hip fracture between the baseline and five-year follow-up CT exams. We excluded nine of these participants because their baseline CT scan was unsuitable for FEA, resulting in a total of 171 hip-fracture cases for statistical analysis. A total of 877 controls without hip fracture were drawn from the aforementioned random selection of 897 AGES-Reykjavik participants who had both baseline and follow-up CT scans; the remaining 20 subjects were excluded because their baseline hip CT scan was unsuitable for FEA or because we could not ascertain their hip fracture status during the five-year follow-up period. The final hip arm comprised 608 women and 440 men, all Caucasian, the age at baseline spanning 66–93 years.

### Vertebral Fracture Adjudication

Incident vertebral fractures were adjudicated by analyzing the lateral scout of the baseline and five-year follow-up CT exams. First, the five-year follow-up CT scout for all participants were examined, using Genant's Semi-Quantitative technique<sup>(25)</sup> to grade fractures as mild (SQ1), moderate (SQ2), or severe (SQ3). Then, if a fracture was found in any vertebra from T6–L4, the baseline CT scout was read to classify the fracture as prevalent or incident. A fracture was classified as incident if the grade of the vertebra increased from baseline to follow-up, and was otherwise classified as prevalent. The fracture grade for a participant was defined as the highest grade of any incident fracture at any vertebral level at the end of the five-year observation period. The inter-observer agreement (Kappa score) between the study reader and both a radiologist at the Icelandic Heart Association (IHA) ( $\kappa = 0.64$ ) and a physician at Synarc, Inc. ( $\kappa = 0.54$ ) for 38 participants and was deemed to be acceptable.

### CT Imaging

All CT scans of the lumbar spine and proximal femur were acquired at IHA on a single scanner (Siemens Sensation 4) using 120 kVp, a pitch of 1.0, and a modulated tube current with a reference exposure of 150 mAs. Images were reconstructed using a 1.0 mm slice thickness, a B30s soft tissue kernel, and a 50 cm field of view. The same external bone mineral phantom (Image Analysis, Columbia, KY) was imaged in every scan.

### Finite Element Analysis

All image processing and finite element analyses of the baseline CT scans were performed at IHA, by a single analyst, blinded to fracture status, using the VirtuOst<sup>®</sup> software application (O.N. Diagnostics, Berkeley, CA). Construction of the finite element models (Figure 2) is described elsewhere.<sup>(4, 14)</sup> In short, the CT scans were calibrated to convert voxel values to BMD, the bone was segmented, and then resampled into isotropic voxels

(1.0 mm for the spine; 1.5 mm for the hip). Each voxel was then converted into a finite element, and assigned elastic and failure material properties based on empirical relations to BMD.<sup>(4, 26)</sup>

In the spine, displacement boundary conditions were applied to simulate a uniform axial compression on the vertebra, applied through a virtual layer of bone cement, and the vertebral strength was defined as the force at 2% deformation. To calculate a load-to-strength ratio, we used body-weight and height data for each participant to estimate a lumbar compressive force for forward bending at the waist while holding a 10-kg weight.<sup>(27)</sup> The assumed moment arm of the paraspinal muscles opposing forward bending was larger for men (5.86 cm) than for women (5.48 cm).<sup>(28)</sup> Vertebral measurements were taken as the average from analysis of L1 and L2 (84% of analyses), or the average of two other vertebra (14%) or from just a single vertebrae (2%) from T12 through L4 if L1 and L2 were not both suitable for analysis.

In the femur, displacement boundary conditions were applied to simulate a sideways fall with the diaphysis angled at 15° to the ground and at 15° of internal rotation. Femoral strength was defined as the force at 4% deformation. To calculate a load-to-strength ratio, we used body-weight and height data for each participant to estimate the impact force for a sideways fall,<sup>(29)</sup> and we assumed a constant trochanteric soft tissue thickness of 25 mm for all participants. The left femur was analyzed except in three participants due to image artifact or abnormal morphology of the left femur.

### BMD Assessment

In addition to these finite element computations, we also used the VirtuOst software application to measure BMD from the same bones for which strength was measured. For the spine, volumetric BMD of the trabecular bone (vBMD, in mg/cm<sup>3</sup>), compatible with the UCSF reference database,<sup>(21, 30)</sup> was measured using an anteriorly-placed elliptical region of interest (Figure 3). As DXA was not acquired for this study, we used the CT scans to measure a Hologic-equivalent areal BMD (in g/cm<sup>2</sup>) at the femoral neck and total hip regions, and associated T-scores, using a similar approach as reported elsewhere (Figure 3).<sup>(15, 22)</sup> Prospective analysis of an independent cohort of 75 women and men confirmed that our CT-measured areal BMD values were highly correlated with DXA-measured values ( $R^2 = 0.90$  femoral neck areal BMD,  $R^2 = 0.93$  total hip areal BMD; Appendix A).

### Interventional Thresholds for Bone Strength

With clinical translation in mind, interventional thresholds for bone strength were established on an independent cohort before this analysis and were designed to coincide generally with the BMD-based interventional thresholds for osteoporosis and low bone mass (aka osteopenia). These strength thresholds were developed by analyzing 1459 predominantly white women and men in a variety of prior research studies, excluding the AGES-Reykjavik study. The spine analysis included 892 women age 21–89 and 286 men age 22–89, and the hip analysis included 856 women age 21–92 and 336 men age 22–89. To minimize any bias from the use of a particular CT machine, data were pooled from 77 different CT scanners, including GE Medical Systems, Siemens, Philips, and Toshiba models. Strength and BMD measures were obtained from these scans as described above. For the spine, linear regression between vertebral strength and vertebral trabecular BMD was then used to find the sex-dependent strength values corresponding to 80 and 120 mg/cm<sup>3</sup>, the recommended vBMD thresholds for osteoporosis and low bone mass, respectively.<sup>(31, 32)</sup> Similarly for the hip, linear regression between femoral strength and femoral neck areal BMD T-score (using the NHANES III database for T-scores<sup>(33)</sup>) was used to find the sex-dependent strength values corresponding to  $T=-2.5$  and  $T=-1.0$ , the T-

score thresholds for defining osteoporosis and low bone mass, respectively.<sup>(34, 35)</sup> The strength interventional thresholds — “fragile bone strength” corresponding to osteoporosis and “low bone strength” corresponding to low bone mass — were defined as the strength values resulting from these regression analyses, rounded up to the nearest 500 N (Table 1). As per W.H.O. guidelines,<sup>(36)</sup> we used female young-reference values in our BMD T-score calculations for both sexes. All of these reported outcomes, including the interventional thresholds, have been cleared by the FDA for clinical use.

### Statistical Analysis

Statistical analyses were performed using SAS (SAS Institute, Inc., Cary, NC) and STATA (StataCorp LP, College Station, TX). Baseline characteristics were compared between fracture cases and controls using an age-adjusted F-test. To quantify the association between each predictor variable and fracture status, odds ratios and 95% Wald confidence intervals were calculated from logistic regression models with parameters estimated using the maximum likelihood approach (LOGISTIC procedure in SAS). To provide insight into how this association might compare across the sexes in the general population, for each predictor the odds ratio was expressed as a ratio with respect to the standard deviation of the controls pooled across the sexes. Prediction capacity was assessed by analysis of the receiver-operator-characteristic curve, adjusting for age. The spine statistics were further adjusted for body-mass index (BMI) and prevalent fracture. Because the clinical utility in identifying mild vertebral fractures is unclear,<sup>(18, 37)</sup> we performed two separate analyses for incident vertebral fractures: one for all SQ1, SQ2, and SQ3 fractures (n = 167 cases), and one only for moderate/severe (SQ2/SQ3) fractures (n = 96 cases). In the latter, SQ1 fractures were excluded from the analysis, not reassigned as controls. Multi-variable logistic regression models were constructed to determine if strength and the load-to-strength ratio remained associated with fracture independent of age, BMI and BMD. We also calculated the net reclassification improvement (NRI)<sup>(38)</sup> to compare a fracture-risk classification using just BMD alone versus a classification using BMD combined with FEA measures of strength and load-to strength ratio. Finally, to assess if individuals with fragile bone strength were indeed at high risk of fracture, we compared the probability of fracture for strength and BMD, using age-adjusted logistic regression models for an age of 75 years (average for the sample); confidence intervals for the resulting probability curve were calculated using the 95% confidence intervals of the regression-model coefficients.

## RESULTS

For both the spine and hip arms, fracture cases were older than controls (Tables 2 and 3). After accounting for this age difference, the fracture cases still had lower strength and BMD compared to the controls, and higher (worse) values for load-to-strength ratio; vertebral fracture cases were also more likely to have a prior vertebral fracture ( $p < 0.001$ ).

### Vertebral Fracture Predictors

In the spine arm, strength, load-to-strength ratio, and vBMD were highly significantly associated with any incident vertebral fracture (SQ1–3) in both women and men both before and after adjusting for age, BMI and prior vertebral fracture (Table 4). Odds ratios (per unit standard deviation) were generally numerically higher for strength than for vBMD for both women and men. Logistic regression results showed that for men only, both strength ( $p = 0.01$ ) and load-to-strength ratio ( $p = 0.03$ ) remained associated with fracture independent of vBMD. The age-adjusted AUC values showed trends that were similar to those for the odds ratios, values ranging from 0.67–0.70 for women and 0.68–0.71 for men.



After excluding mild (SQ1) fractures from analysis, the age-adjusted odds ratios of all predictor variables increased for women (Table 4). For men, the odds ratios for strength and the load-to-strength ratio were generally unchanged, while the odds ratio for vBMD was no longer significant. After further adjusting for BMI and prior vertebral fracture, all three predictors remained significant for women, but none remained significant for men, although the small number of cases for men ( $n = 21$ ) compromised statistical power for this analysis.

### Hip Fracture Predictors

In the hip arm, strength, load-to-strength ratio, femoral neck areal BMD, and total hip areal BMD were all highly significantly associated with incident hip fractures in both women and men, both before and after adjusting for age and BMI (Table 5). Further, after adjusting for age, BMI, and femoral neck areal BMD using logistic regression models, both strength ( $p = 0.01$ ) and the load-to-strength ratio ( $p = 0.005$ ) remained associated with fracture for women, but not for men. When total hip areal BMD was placed in the multivariate model instead of femoral neck areal BMD, each of strength (women  $p = 0.0006$ , men  $p = 0.0001$ ) and load-to-strength ratio (women  $p = 0.015$ , men  $p = 0.010$ ) remained associated with fracture for both sexes. The age-adjusted AUC values tended to be higher for men (0.84–0.86) than women (0.78–0.80), but otherwise showed trends that were similar to those for the odds ratios.

### Reclassification Analysis

Consistent with the trends for odds ratios, the reclassification analysis revealed an added benefit of combining strength and BMD compared to the use of BMD alone. For the spine arm, in men, the combination of strength and vBMD (NRI = 62%,  $p = 0.006$ ), or strength, load-to-strength ratio, and vBMD (NRI = 63%,  $p = 0.005$ ) significantly improved fracture classification for moderate/severe (SQ2/3) vertebral fractures; this effect was not significant when all fractures (SQ1/2/3) were considered. No significant improvement was seen for women ( $p > 0.5$ ). For the hip arm, combining femoral strength and femoral neck areal BMD improved classification of hip fractures (NRI = 33%,  $p = 0.002$ ) in women, as did combining strength, the load-to-strength ratio, and femoral neck areal BMD (NRI = 37%,  $p = 0.001$ ). No significant improvement was seen for men ( $p > 0.4$ ).

### Interventional Thresholds

The age-adjusted logistic regression analysis indicated that, at an average age of 75 years, the calculated probability of fracture for this analysis sample was similarly high at the BMD and bone strength interventional thresholds, both for women and men and at the spine (Figure 4) and the hip (Figure 5). For example, the probability of vertebral fracture at the thresholds for fragile bone strength and osteoporosis were 22.2% (95% CI: 18.5–26.4%) and 14.6% (11.1–19.0%), respectively, for the women, and 14.8% (11.2–19.3%) and 14.6% (11.1–19.0%), respectively, for the men. At the hip, the probability of hip fracture associated with the thresholds for fragile bone strength and osteoporosis were 17.1% (13.5 – 21.3%) and 21.8% (17.0 – 27.5%), respectively, for the women, and 24.0% (17.0 – 32.7%) and 33.4% (23.3 – 45.4%), respectively, for the men.

## DISCUSSION

The finite element analysis technique, which uses computational biomechanical principles coupled with patient-specific information in clinical CT scans to mechanistically simulate bone failure, has been validated in cadaver studies by numerous groups for both the spine<sup>(4–6, 39)</sup> and hip<sup>(7–12)</sup> and clinically has been shown to be significantly associated with incident and prevalent fracture in multiple cohorts.<sup>(13–18, 20, 40)</sup> Our new data provide further clinical validation. For the spine, vertebral strength was associated with fracture in both

women and men; consistently had the (numerically) highest odds ratios compared to the other predictors; was associated with fracture independently of vBMD in men; and for men with more severe fracture (SQ2/3), remained significant whereas vBMD lost significance. For the hip, femoral strength was associated with fracture independently of femoral neck areal BMD in women and total hip areal BMD in both sexes. With clinical translation in mind, we introduced and prospectively evaluated interventional thresholds for bone strength, and confirmed that these thresholds for “fragile bone strength” were associated with fracture probability levels equivalent to those for well-established thresholds for osteoporosis, both at the hip and spine and in women and men. The probability of fracture in this study depends on the nature of the case-control design and does not represent actual clinical fracture risk. However, since it is well established that individuals over age 65 who have osteoporosis at the hip or spine by BMD criteria should be considered at high clinical risk of fracture, and since strength was associated with fracture at least as well as was BMD at both the hip and spine in the present study, these results indicate that individuals who have “fragile bone strength” at the hip or spine should also be considered at high clinical risk of fracture.

One novel aspect of this study is the use of FEA-based vertebral strength assessment, and vertebral trabecular BMD, for prediction of incident vertebral fractures in women, the first study of its kind. Our findings are consistent with those from the Osteoporotic Fractures in Men (MrOS) study of men over age 65, which showed that vertebral compressive strength (and the load-to-strength ratio) were highly significant predictors of incident clinical vertebral fracture in men.<sup>(13)</sup> Similar to our reported odds ratios, hazard ratios for strength in that study were numerically higher for strength compared to vBMD, although that study evaluated integral vBMD (trabecular and cortical) and not trabecular vBMD as in this study. We are aware of no studies for women reporting on incident spine fracture and CT-based BMD or strength. In the current study, the association between vertebral strength and vertebral fracture was uniformly stronger in women than men, both before and after adjusting for age and prevalent fracture, and in particular when restricted to more severe fractures (SQ2/3). Spine DXA was not used in this study. However, in the MrOS study of elderly men, the age-adjusted hazard ratio was two-fold higher ( $p < 0.01$ ) for strength than for DXA-measured lumbar spine areal BMD, and strength was associated with fracture independent of DXA areal BMD.

Although no prior study has investigated FEA predictors of incident vertebral fracture in women, several cross-sectional studies in the U.S. and Japan have shown statistically significant associations between FEA and prevalent vertebral fracture, supporting the generality of our spine results. In 1991, Faulkner found that an FEA-estimated vertebral yield stress better distinguished women aged 20–79 years with a prevalent vertebral fracture than did a QCT-based measure of bone mineral content.<sup>(16)</sup> More recently, two studies of women over age 50 in the Rochester MN area found that vertebral strength, load-to-strength ratio, and vBMD were all highly associated with prevalent vertebral fractures.<sup>(17, 18)</sup> Imai *et al.*<sup>(40)</sup> also found a trend for the association of prevalent vertebral fracture in Japanese women to be greater for strength than for CT-measured vBMD or DXA-measured spine areal BMD. In all these more recent studies, vertebral strength consistently performed statistically better than lumbar spine areal BMD as measured by DXA. DXA at the spine is limited by its two-dimensional nature and its inclusion of the posterior elements and any aortic calcification in the BMD measure. Given these limitations, and since our findings show consistent agreement with these prior FEA studies, the collective literature suggests that vertebral strength can reasonably be expected to perform better than DXA for assessing risk of incident spine fractures in both women and men.

Another novel aspect of this study was our prospective validation of previously established interventional thresholds for bone strength, which enables FEA to be used clinically to

identify women and men at high risk of fracture. We found that the women and men in this study who had “fragile bone strength” were at an equivalently high probability of fracture as were the women and men who had BMD-defined osteoporosis, the latter criterion placing them clinically in a “high-risk” category for fracture. Since the thresholds for bone strength were derived from a previously measured strength-BMD relationship for Caucasians, these thresholds should remain valid for any population with a similar strength-BMD relationship. The AGES-Reykjavik cohort shows such a correspondence (Figure 6), the value of strength at the threshold for fragile bone strength being — as per design — just slightly higher than the value of strength directly corresponding to the osteoporosis threshold. In general, the relation between whole-bone strength and BMD by FEA analysis depends on such morphological characteristics as the size and shape of the bone and the spatial distribution of bone density, including the trabecular-cortical characteristics. As such, the strength thresholds reported here might not be directly applicable to bones in non-Caucasian populations having different morphological characteristics that would alter the relation between BMD and whole-bone strength.

For the hip, our findings for women that femoral strength was associated with fracture independently of areal BMD, and that reclassification improved when using a combination of femoral strength and areal BMD, together suggest that more individuals at high risk of fracture can be identified by using measures of both femoral strength and hip BMD than by using measures of hip BMD alone. Part of this effect is that some individuals with low bone mass (aka osteopenia) who fractured also had “fragile bone strength,” as shown in a plot of femoral strength and femoral neck areal BMD from this study (Figure 6, see shaded region), and as observed also in the MrOS study of elderly men.<sup>(14)</sup> The underlying biophysical mechanisms for this effect are not yet clear, perhaps related to geometry, or relatively low trabecular to cortical mass,<sup>(41)</sup> or locally weak regions within the bone, any of which might go undetected by an areal BMD measure due to its projectional nature. Regardless, these findings illustrate that women and men who have low bone mass can be at as high a risk of fracture as the risk associated with having osteoporosis if they also have “fragile bone strength.” Whether or not such osteopenic high-risk individuals would correspond with those identified using an absolute-risk approach that incorporates various clinical risk factors<sup>(42, 43)</sup> is unclear, and remains a topic for future research.

The generality of our hip strength results is supported by reports of similar findings from the only other two incident hip-fracture studies that analyzed both areal BMD and FEA-estimated hip strength. Femoral neck areal BMD was not reported in either of these two studies but total hip areal BMD was. In the first prior study, an analysis of incident hip fractures in men over age 65 in the MrOS cohort<sup>(14)</sup> that used the same software as in the current analysis, age-adjusted hazard ratios were numerically higher for femoral strength (6.5, 95% CI: 2.3–18.3) than for DXA-measured total hip areal BMD (4.4, 95% CI: 2.1–9.1). That finding is consistent with ours of a higher odds ratio for femoral strength compared to total hip areal BMD. In the second prior study, an age- and sex-matched nested case-control analysis of a subset of the AGES-Reykjavik participants performed using different image-processing and FEA software by Keyak *et al.*,<sup>(15)</sup> femoral strength in a stance loading configuration remained a significant predictor of hip fracture after accounting for total hip areal BMD in men ( $p = 0.01$ ), and just missed statistical significance for women ( $p = 0.06$ ). Similar trends ( $p = 0.06$ ) were seen in the fall loading condition for both women and men, and it is likely these trends would have reached statistical significance had the number of fractures in that analysis (71 women and 45 men) been greater. These results are therefore also consistent with our findings.

Despite this consistency between these past studies and the current study, Keyak *et al.* concluded that femoral strength may be a more important fracture risk predictor for men



than for women, in apparent contradiction to our findings that the odds ratio for femoral strength was higher for women than for men. The Keyak *et al.* conclusion was based on their finding that the ratio of the mean difference in femoral strength between their age-matched cases and controls, divided by the sex-specific standard deviation, was larger for men (ratio = 0.72 for men vs. 0.32 for women). We used a sex-pooled standard deviation in our odds ratio calculations. However, to compare against Keyak *et al.*, we also normalized our strength differences by sex-specific standard deviations and then adjusted our strength results to age 80 (the mean age in the Keyak *et al.* study) and found a similar trend of a higher ratio for men than women (ratio = 0.83 for men vs. 0.59 for women), indicating congruence in the two studies. Further, when we also normalized our logistic regression parameters by sex-specific standard deviations, we found that the age-adjusted odds ratio (95% CI) for strength was numerically higher for men 3.2 (2.1 – 4.7) than women 2.8 (2.0 – 3.9), again consistent with the Keyak *et al.* findings. This apparent reversal in our odds ratios was due to the women having a lower standard deviation relative to the sex-pooled standard deviation than men, which in turn was due to the lower mean value of femoral strength for women. For our primary analysis, we normalized by the sex-pooled standard deviation in order to provide insight into risk differences between women and men. Our finding of a higher odds ratio for women than men, when using a sex-pooled standard deviation, indicates that a fixed decrement of bone strength elevates risk more for women than for men. It is well established that women lose femoral strength at a greater absolute rate with aging than do men.<sup>(44, 45)</sup> Thus, our results help explain the known higher rate of hip fracture in women than men. Further, evaluated in this way, our results suggest that the association between femoral strength and hip fracture is at least as important for women as for men.

There are a number of limitations for this study. Most importantly, the analysis sample lacked racial variation, and as noted above, potential differences in the general relation between areal BMD and whole-bone strength may affect the generality of the strength thresholds in certain non-Caucasian populations. However, since hip strength and hip areal BMD are quite well correlated (see Figure 6, for example), and since hip areal BMD is a robust predictor of fracture across races,<sup>(36)</sup> strength should also be associated with fracture risk across races. This is consistent with results from prevalent and incident fracture-outcome studies from multiple different cohorts from the U.S.,<sup>(13, 14, 17, 18)</sup> Japan<sup>(40)</sup> and Iceland.<sup>(15)</sup> A second limitation is that we did not use DXA. However, as shown by others<sup>(15, 22)</sup> and by our own data (Appendix A), the hip areal BMD T-score as measured from CT is highly correlated and numerically equivalent to that as measured by DXA, and thus our results should remain substantially unchanged had real DXA been used. Even so, future studies are required to confirm the expected advantage of vertebral strength over DXA for predicting incident spine fractures in women, and more generally to confirm our various results in other cohorts.

An additional limitation is that we used a case-control approach rather than a case-cohort approach that would have allowed direct estimation of prevalence and absolute risk. This choice was a trade-off in the spine arm to increase statistical power by including morphologic vertebral fractures as (incident) cases, which required follow-up CT scans for adjudication. Excluding those without a follow-up exam meant that the random sample, while useful for selecting controls, under-represented cases due to an association between incident fracture and drop-out. Even so, the case-control design still allowed us to evaluate odds ratios for strength and to show that the probability of fracture associated with fragile bone strength and osteoporosis were similarly high. We note also that despite the general trend for larger odds ratios for strength compared to BMD, there were only small differences in the AUC values between the various predictors. This is not surprising, since typically large differences in an odds ratio need to occur before the AUC changes appreciably.<sup>(46)</sup>

However, since the AUC represents the performance of a predictor across the entire range of sensitivities and specificities, a finding of only a small difference in AUC values between predictors may not represent potential benefits in clinical practice and decision-making, which depends more on where an individual patient falls with respect to any relevant interventional threshold. For example, we found improved reclassification when both strength and BMD were used instead of just BMD, presumably because a statistically significant subset of individuals had lower than expected bone strength for their BMD and were therefore at higher than expected risk of fracture. Clinically, that should translate to a subset of individuals with low bone mass, but fragile bone strength, who are indeed at higher risk of fracture.

There are also inherent limitations with the finite element technique. While our technique has shown good agreement in strength values compared to cadaver testing,<sup>(4, 8)</sup> our current implementation did not include some potentially important features, such as micro-scale effects, fine resolution of the thin cortex, and multiple loading conditions. There is also potential to improve the load-to-strength ratio formulation, which in this study did not include patient-specific modeling of the intervertebral discs, spinal curvature, or muscle morphology. In the hip, a CT-based measure of patient-specific soft-tissue thickness over the greater trochanter was not included, although our preliminary analyses indicated that including such detail did not improve the age-adjusted odds ratio after adjustment for BMI. Development of methods to better utilize such patient-specific model inputs remains a topic for future research.

When viewed in the context of the available literature that has now accumulated on FEA of CT scans, these new results suggest that FEA (which clinically would also include a CT-based BMD analysis) can provide an alternative clinical tool to DXA for meeting the increasing need for additional osteoporosis and fracture risk assessment.<sup>(3)</sup> While the use of a dedicated CT scan would introduce additional radiation compared to the use of DXA, the use of an ancillary approach,<sup>(47)</sup> in which a previously acquired CT is utilized,<sup>(48–50)</sup> would circumvent this limitation, and tens of millions of such CT scans are taken annually in women and men over age 50.<sup>(51)</sup> As noted above, such an approach may not only identify individuals with osteoporosis and fragile bone strength, but also a subset of individuals with low bone mass who have fragile bone strength and are thus at high risk of fracture.

## Acknowledgments

Funding for this research was provided by NIH AR052234, NIH AR057616 and the Intramural Research Program of the NIH, National Institute on Aging. The CT scans used to establish equivalence between CT and DXA-measured areal BMD were acquired as part of the Rochester Epidemiology Project at Mayo Clinic Rochester (PI: Dr. Sundeep Khosla) and used here with permission. Verification of the vertebral fracture adjudication procedure at Synarc, Inc. was kindly arranged by Dr. Thomas Fuerst.

## References

1. 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]
2. 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]
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. Crawford RP, Cann CE, Keaveny TM. Finite element models predict in vitro vertebral body compressive strength better than quantitative computed tomography. *Bone*. 2003; 33(4):744–50. [PubMed: 14555280]

5. Imai K, Ohnishi I, Yamamoto S, Nakamura K. In vivo assessment of lumbar vertebral strength in elderly women using computed tomography-based nonlinear finite element model. *Spine*. 2008; 33(1):27–32. [PubMed: 18165745]
6. Dall'Ara E, Pahr D, Varga P, Kainberger F, Zysset P. QCT-based finite element models predict human vertebral strength in vitro significantly better than simulated DEXA. *Osteoporos Int*. 2012; 23(2):563–72. [PubMed: 21344244]
7. Keyak JH. Improved prediction of proximal femoral fracture load using nonlinear finite element models. *Med Eng Phys*. 2001; 23(3):165–73. [PubMed: 11410381]
8. Roberts BJ, Kopperdahl DL, Thrall E, Muller JA, Keaveny TM, Bouxsein ML. Prediction of femoral strength in a sideways fall configuration using QCT-based finite element analysis. *Bone*. 2009; 44(S1):S72.
9. Bessho M, Ohnishi I, Matsuyama J, Matsumoto T, Imai K, Nakamura K. Prediction of strength and strain of the proximal femur by a CT-based finite element method. *J Biomech*. 2007; 40(8):1745–53. [PubMed: 17034798]
10. Schileo E, Taddei F, Cristofolini L, Viceconti M. 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*. 2008; 41(2):356–67. [PubMed: 18022179]
11. Koivumaki JE, Thevenot J, Pulkkinen P, Kuhn V, Link TM, Eckstein F, Jamsa T. Ct-based finite element models can be used to estimate experimentally measured failure loads in the proximal femur. *Bone*. 2012; 50(4):824–9. [PubMed: 22306697]
12. Dall'Ara E, Luisier B, Schmidt R, Kainberger F, Zysset P, Pahr D. A nonlinear QCT-based finite element model validation study for the human femur tested in two configurations in vitro. *Bone*. 2013; 52(1):27–38. [PubMed: 22985891]
13. Wang X, Sanyal A, Cawthon PM, Palermo L, Jekir M, Christensen J, Ensrud KE, Cummings SR, Orwoll E, Black DM, Keaveny TM. Prediction of new clinical vertebral fractures in elderly men using finite element analysis of CT scans. *J Bone Miner Res*. 2012; 27(4):808–16. [PubMed: 22190331]
14. Orwoll ES, Marshall LM, Nielson CM, Cummings SR, Lapidus J, Cauley JA, Ensrud K, Lane N, Hoffmann PR, Kopperdahl DL, Keaveny TM. 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]
15. Keyak JH, Sigurdsson S, Karlsdottir G, Oskarsdottir D, Sigmarsdottir A, Zhao S, Kornak J, Harris TB, Sigurdsson G, Jonsson BY, Siggeirsdottir K, Eiriksdottir G, Gudnason V, Lang TF. 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]
16. Faulkner KG, Cann CE, Hasegawa BH. Effect of bone distribution on vertebral strength: assessment with patient-specific nonlinear finite element analysis. *Radiology*. 1991; 179(3):669–74. [PubMed: 2027972]
17. Melton LJ, Riggs BL, Keaveny TM, Achenbach SJ, Hoffmann PF, Camp JJ, Rouleau PA, Bouxsein ML, Amin S, Atkinson EJ, Robb RA, Khosla S. Structural determinants of vertebral fracture risk. *J Bone Miner Res*. 2007; 22(12):1885–92. [PubMed: 17680721]
18. 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. Relation of vertebral deformities to bone density, structure, and strength. *J Bone Miner Res*. 2010; 25(9):1922–30. [PubMed: 20533526]
19. Graeff C, Marin F, Petto H, Kayser O, Reisinger A, Pena J, Zysset P, Gluer CC. High resolution quantitative computed tomography-based assessment of trabecular microstructure and strength estimates by finite-element analysis of the spine, but not DXA, reflects vertebral fracture status in men with glucocorticoid-induced osteoporosis. *Bone*. 2012; 52(2):568–77. [PubMed: 23149277]
20. Amin S, Kopperdahl DL, Melton LJ 3rd, Achenbach SJ, Therneau TM, Riggs BL, Keaveny TM, Khosla S. 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]
21. Block JE, Smith R, Gluer CC, Steiger P, Ettinger B, Genant HK. Models of spinal trabecular bone loss as determined by quantitative computed tomography. *J Bone Miner Res*. 1989; 4(2):249–57. [PubMed: 2728928]

22. Khoo BC, Brown K, Cann C, Zhu K, Henzell S, Low V, Gustafsson S, Price RI, Prince RL. Comparison of QCT-derived and DXA-derived areal bone mineral density and T scores. *Osteoporos Int.* 2009; 20(9):1539–45. [PubMed: 19107384]
23. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF, Hoffman HJ, Gudnason V. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol.* 2007; 165(9):1076–87. [PubMed: 17351290]
24. Siggeirsdottir K, Aspelund T, Sigurdsson G, Mogensen B, Chang M, Jonsdottir B, Eiriksdottir G, Launer LJ, Harris TB, Jonsson BY, Gudnason V. Inaccuracy in self-report of fractures may underestimate association with health outcomes when compared with medical record based fracture registry. *Eur J Epidemiol.* 2007; 22(9):631–9. [PubMed: 17653601]
25. Genant HK, Wu CY, Vankuijk C, Nevitt MC. Vertebral fracture assessment using a semiquantitative technique. *J Bone Miner Res.* 1993; 8(9):1137–48. [PubMed: 8237484]
26. Kopperdahl DL, Morgan EF, Keaveny TM. Quantitative computed tomography estimates of the mechanical properties of human vertebral trabecular bone. *J Orthop Res.* 2002; 20(4):801–5. [PubMed: 12168670]
27. Bouxsein ML, Melton LJ 3rd, Riggs BL, Muller J, Atkinson EJ, Oberg AL, Robb RA, Camp JJ, Rouleau PA, McCollough CH, Khosla S. Age- and sex-specific differences in the factor of risk for vertebral fracture: a population-based study using QCT. *J Bone Miner Res.* 2006; 21(9):1475–82. [PubMed: 16939406]
28. Wilson, S. Development of a Model to Predict the Compressive Forces on the Spine Associated With Age-Related Vertebral Fractures [dissertation]. Cambridge (MA): Massachusetts Institute of Technology; 1994.
29. van den Kroonenberg AJ, Hayes WC, McMahon TA. Dynamic models for sideways falls from standing height. *J Biomech Eng.* 1995; 117(3):309–18. [PubMed: 8618384]
30. Cann CE, Genant HK, Kolb FO, Ettinger B. Quantitative computed tomography for prediction of vertebral fracture risk. *Bone.* 1985; 6:1–7. [PubMed: 3994856]
31. Engelke K, Adams JE, Armbrrecht G, Augat P, Bogado CE, Bouxsein ML, Felsenberg D, Ito M, Prevrhal S, Hans DB, Lewiecki EM. 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.* 2008; 11(1):123–62. [PubMed: 18442757]
32. ACR Practice Guideline for the Performance of Quantitative Computed Tomography (QCT) Bone Densitometry. American College of Radiology; 2008.
33. Looker AC, Wahner HW, Dunn WL, Calvo MS, Harris TB, Heyse SP, Johnston CC Jr, Lindsay R. Updated data on proximal femur bone mineral levels of US adults. *Osteoporos Int.* 1998; 8(5): 468–89. [PubMed: 9850356]
34. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. World Health Organization technical report series. 1994; 843:1–129. [PubMed: 7941614]
35. Lewiecki EM, Gordon CM, Baim S, Leonard MB, Bishop NJ, Bianchi ML, Kalkwarf HJ, Langman CB, Plotkin H, Rauch F, Zemel BS, Binkley N, Bilezikian JP, Kendler DL, Hans DB, Silverman S. International Society for Clinical Densitometry 2007 Adult and Pediatric Official Positions. *Bone.* 2008; 43(6):1115–21. [PubMed: 18793764]
36. Kanis JA, McCloskey EV, Johansson H, Oden A, Melton LJ 3rd, Khaltav N. A reference standard for the description of osteoporosis. *Bone.* 2008; 42(3):467–75. [PubMed: 18180210]
37. Ferrar L, Jiang G, Cawthon PM, San Valentin R, Fullman R, Lambert L, Cummings SR, Black DM, Orwoll E, Barrett-Connor E, Ensrud K, Fink HA, Eastell R. Identification of vertebral fracture and non-osteoporotic short vertebral height in men: the MrOS study. *J Bone Miner Res.* 2007; 22(9):1434–41. [PubMed: 17563237]
38. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med.* 2008; 27(2):157–72. discussion 207–12. [PubMed: 17569110]
39. Buckley JM, Loo K, Motherway J. Comparison of quantitative computed tomography-based measures in predicting vertebral compressive strength. *Bone.* 2007; 40(3):767–74. [PubMed: 17174619]

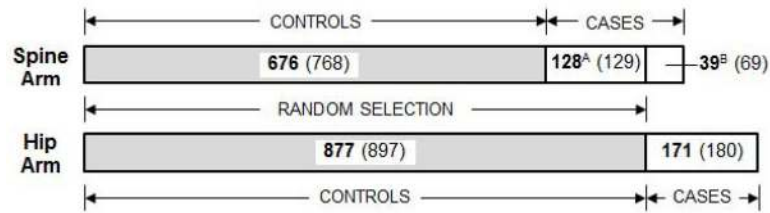
40. Imai K, Ohnishi I, Matsumoto T, Yamamoto S, Nakamura K. Assessment of vertebral fracture risk and therapeutic effects of alendronate in postmenopausal women using a quantitative computed tomography-based nonlinear finite element method. *Osteoporos Int.* 2009; 20(5):801–10. [PubMed: 18800178]
41. Nawathe S, Akhlaghpour H, Bouxsein ML, Keaveny TM. Microstructural failure mechanisms in the human proximal femur for sideways fall loading. *J Bone Miner Res.* 2013
42. Black DM, Steinbuch M, Palermo L, Dargent-Molina P, Lindsay R, Hoseyni MS, Johnell O. An assessment tool for predicting fracture risk in postmenopausal women. *Osteoporosis Int.* 2001; 12(7):519–28.
43. 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. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. *Osteoporosis Int.* 2007; 18(8):1033–46.
44. Pluijm SM, Koes B, de Laet C, Van Schoor NM, Kuchuk NO, Rivadeneira F, Mackenbach JP, Lips P, Pols HA, Steyerberg EW. A simple risk score for the assessment of absolute fracture risk in general practice based on two longitudinal studies. *J Bone Miner Res.* 2009; 24(5):768–74. [PubMed: 19113932]
45. Keaveny TM, Kopperdahl DL, Melton LJ 3rd, Hoffmann PF, Amin S, Riggs BL, Khosla S. Age-dependence of femoral strength in white women and men. *J Bone Miner Res.* 2010; 25(5):994–1001. [PubMed: 19874201]
46. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *American Journal of Epidemiology.* 2004; 159(9):882–90. [PubMed: 15105181]
47. 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]
48. Mueller DK, Kutscherenko A, Bartel H, Vlassenbroek A, Ourednicek P, Erckenbrecht J. Phantomless QCT BMD system as screening tool for osteoporosis without additional radiation. *Eur J Radiol.* 2011; 79(3):375–81. [PubMed: 20223609]
49. Pickhardt PJ, Lee LJ, del Rio AM, Lauder T, Bruce RJ, Summers RM, Pooler BD, Binkley N. 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]
50. Summers RM, Baecher N, Yao J, Liu J, Pickhardt PJ, Choi JR, Hill S. 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]
51. Berrington de Gonzalez A, Mahesh M, Kim KP, Bhargavan M, Lewis R, Mettler F, Land C. Projected cancer risks from computed tomographic scans performed in the United States in 2007. *Arch Intern Med.* 2009; 169(22):2071–7. [PubMed: 20008689]

## Appendix A: Equivalence between CT and DXA-measured areal BMD

After developing CT-based hip areal BMD algorithms on a first cohort, we then tested those algorithms on a second cohort— all independent of the AGES-Reykjavik participants. The algorithms were calibrated for consistency with the NHANES III reference database. Participants in the validation cohort comprised 75 community-dwelling Caucasian women and men over age 65 from the Rochester, MN area. They were scanned on a Siemens Sensation 64 CT scanner and a Lunar Prodigy DXA scanner. For the femoral neck, mean  $\pm$  standard deviation T-scores were  $-0.9 \pm 1.1$  for DXA and  $-1.0 \pm 1.0$  for CT. While this difference was statistically significant ( $p < 0.0001$ , paired t-test), it was so small as to be clinically insignificant, and there was a high degree of correlation ( $R^2 = 0.90$ ). For the total hip, there was no significant difference in T-scores (DXA =  $-0.3 \pm 1.3$ ; CT =  $-0.2 \pm 1.2$ ;  $p =$

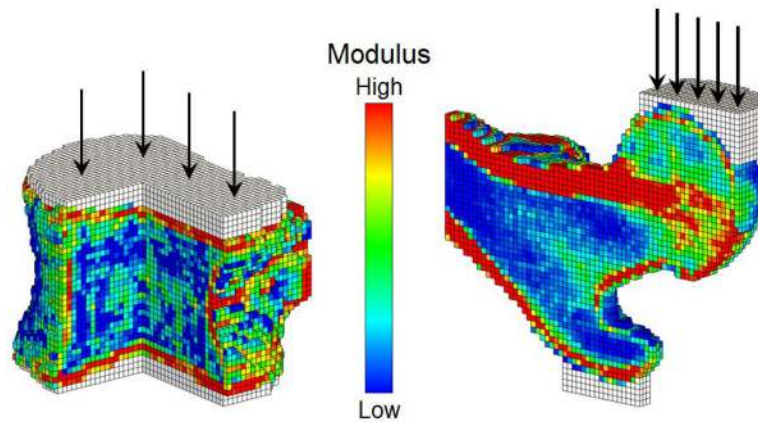


0.19 paired t-test;  $R^2 = 0.93$ ). These results confirmed that our use of CT in this study was a valid substitute for DXA in terms of providing DXA-equivalent T-scores at the hip.

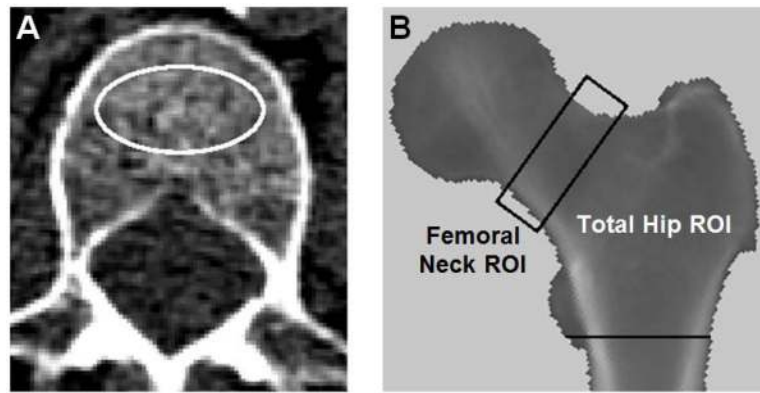


**Figure 1.**

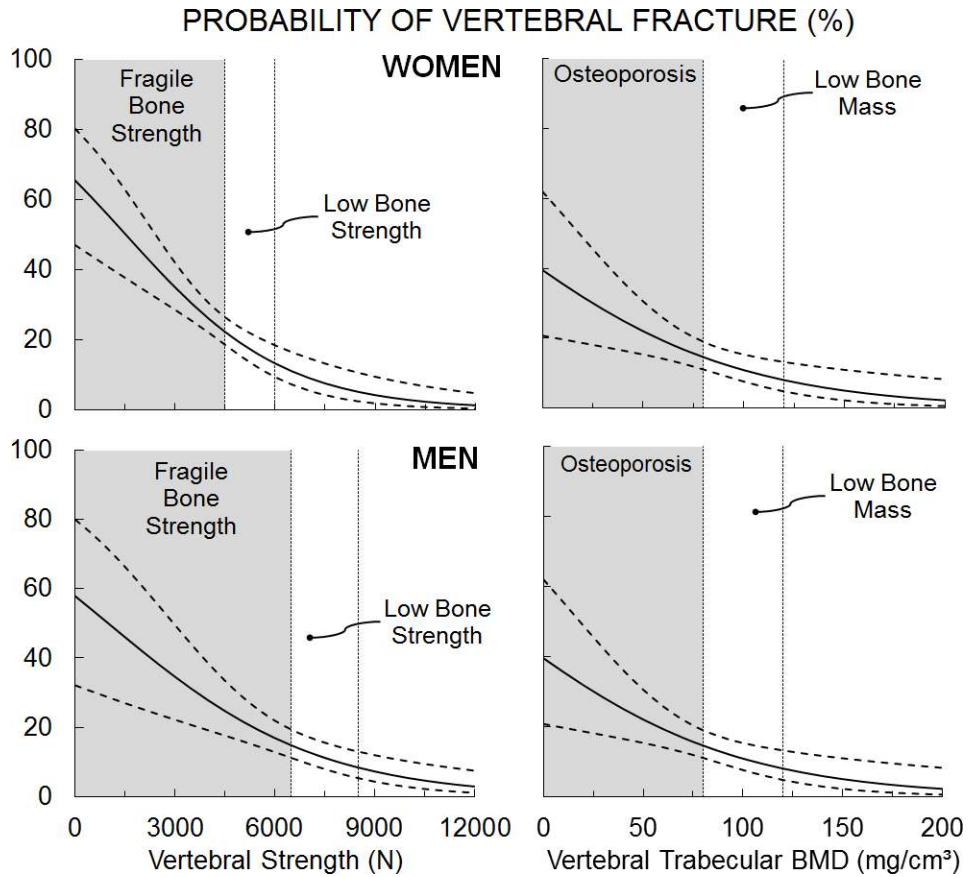
Construction of the case-control analysis sample for the spine arm (top) and hip arm (bottom) showing the sample size after (and before) exclusions (see text for details). The spine arm comprised 128 morphologic (A) and 39 clinical (B) incident vertebral fracture cases, and 676 controls without incident vertebral fracture (controls in gray). The hip arm comprised 171 cases with incident hip fracture and 877 controls without incident hip fracture (controls in gray). The same random selection used in both arms (initially 897 participants) was subject to the requirement of having both a baseline and a follow-up CT scan. Due to an association between clinical fracture and drop-out, this requirement led to a lack in the random selection of participants having clinical vertebral or hip fracture.



**Figure 2.** Sectioned views of finite element models of a vertebral body (left) and a proximal femur (right) showing the distribution of elastic modulus. Applied loads, which simulate axial compression for the spine and an unprotected sideways fall for the hip, are shown schematically, applied through layers of bone cement (white elements) to distribute the load over the bone surface.



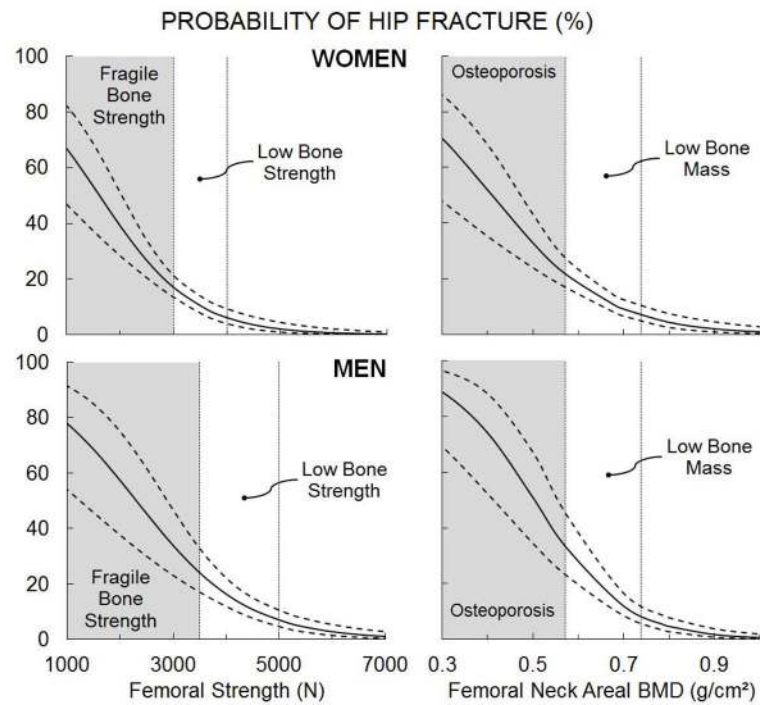
**Figure 3.** Vertebral trabecular BMD was measured from an elliptical region-of-interest covering the middle 9 mm of the vertebral body height (left). Femoral neck and total hip areal BMD were measured from a two-dimensional projection of the re-oriented CT scan (right). The total hip region extended from below the lesser trochanter to the proximal edge of the femoral neck box.



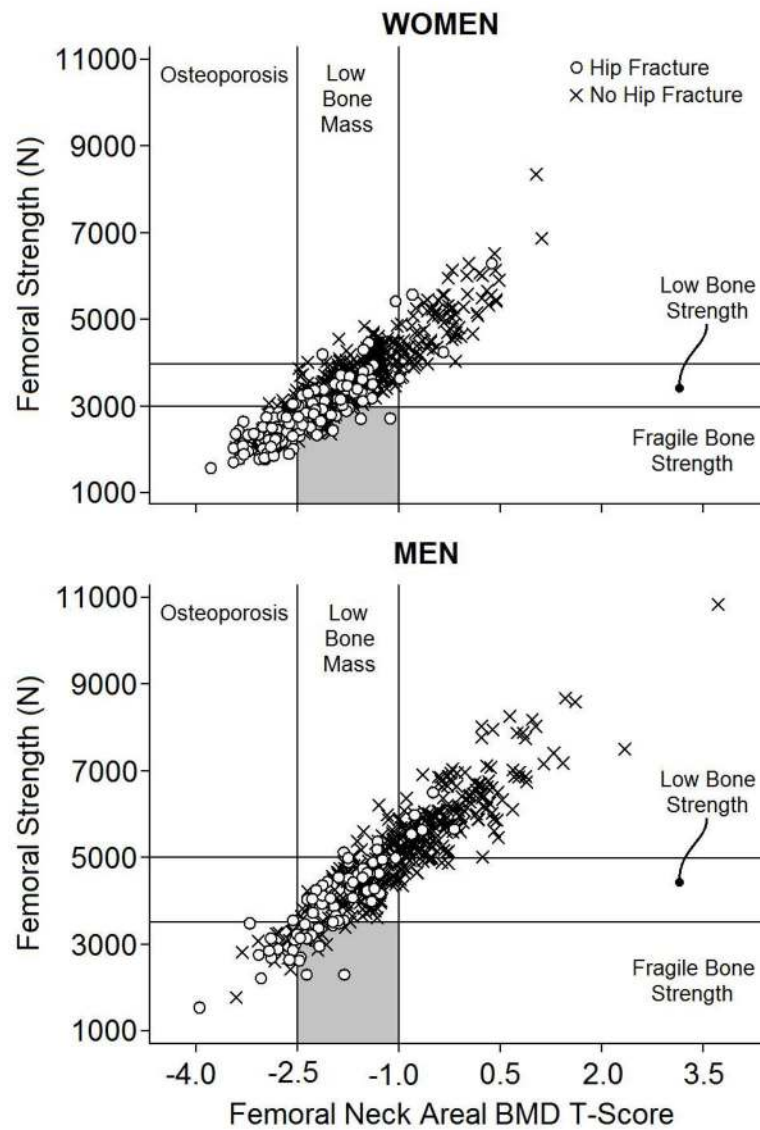
**Figure 4.**

The probability of vertebral fracture at age 75 versus vertebral strength (left) and trabecular BMD (right) for women (top) and men (bottom), using the age-adjusted logistic regression models. The shaded “high-risk” regions are defined by values of strength and BMD below the thresholds for “fragile bone strength” and osteoporosis, respectively (See Table 1 for the bone strength thresholds). Dashed lines show the 95% confidence limits.





**Figure 5.** The probability of hip fracture calculated at age 75 versus femoral strength (left) and femoral neck areal BMD (right) for women (top) and men (bottom), using the age-adjusted logistic regression models. See Figure 4 for details.



**Figure 6.** Femoral strength versus femoral neck areal BMD for women (top) and men (bottom), showing the interventional thresholds for fragile bone strength, low bone strength, osteoporosis, and low bone mass. The shaded regions identify the subset of individuals with low bone mass (aka osteopenia) who also have fragile bone strength.

**Table 1**

Intervention thresholds for “fragile bone strength” and “low bone strength.”

	<b>Fragile Bone Strength <sup>a</sup></b>		<b>Low Bone Strength <sup>b</sup></b>	
	<b>Women</b>	<b>Men</b>	<b>Women</b>	<b>Men</b>
Vertebral Strength (N)	4,500	6,500	6,000	8,500
Femoral Strength (N)	3,000	3,500	4,000	5,000

<sup>a</sup> corresponds to osteoporosis by BMD classification.

<sup>b</sup> corresponds to low bone mass (or osteopenia) by BMD classification.

**Table 2**

Mean (standard deviation) baseline characteristics of the spine arm.

	No-Fracture Controls	SQ1, SQ2, SQ3 Fracture Cases	SQ2, SQ3 Fracture Cases	p-value <sup>a</sup>	Mean (SD)	p-value <sup>a</sup>
<i>Women</i>						
Sample size	380	117	75			
Number with prior vertebral fracture	48	52	22			
Age (years)	74.3 (5.2)	76.4 (5.4)	76.7 (5.0)	<0.001		<0.001
BMI (kg/m <sup>2</sup> )	27.5 (4.4)	26.8 (4.8)	26.7 (4.3)	0.30		0.31
Strength (N)	4880 (1670)	3900 (1320)	3650 (1130)	<0.0001		<0.0001
Load-to-Strength Ratio	0.43 (0.15)	0.51 (0.16)	0.55 (0.17)	<0.0001		<0.0001
Vertebral Trabecular BMD (mg/cm <sup>3</sup> )	82.0 (32.8)	60.3 (26.5)	55.8 (22.6)	<0.0001		<0.0001
<i>Men</i>						
Sample size	296	50	21			
Number with prior vertebral fracture	54	27	5			
Age (years)	74.8 (5.1)	76.7 (5.3)	76.5 (5.5)	<0.05		0.14
BMI (kg/m <sup>2</sup> )	27.0 (3.7)	26.0 (3.6)	25.0 (3.2)	0.18		<0.05
Strength (N)	7540 (2470)	5950 (1880)	5840 (2030)	<0.0001		<0.01
Load-to-Strength Ratio	0.33 (0.11)	0.41 (0.13)	0.40 (0.14)	<0.0001		<0.01
Vertebral Trabecular BMD (mg/cm <sup>3</sup> )	91.6 (32.7)	73.9 (30.5)	79.0 (38.9)	<0.01		0.16

<sup>a</sup> p-values comparing cases versus controls are from a t-test for age, and from an age-adjusted F-test for all other comparisons.

**Table 3**

Mean (standard deviation) baseline characteristics of the hip arm.

	No-Fracture Controls	Hip Fracture Cases	p-value <sup>a</sup>
<i>Women</i>			
Sample size	500	108	
Age (years)	74.7 (5.3)	79.4 (5.7)	< 0.0001
BMI (kg/m <sup>2</sup> )	27.2 (4.5)	25.7 (5.2)	0.054
Strength (N)	3600 (910)	2800 (670)	< 0.0001
Load-to-Strength Ratio	0.95 (0.22)	1.12 (0.25)	< 0.0001
Femoral Neck Areal BMD (g/cm <sup>2</sup> )	0.68 (0.11)	0.59 (0.08)	< 0.0001
Total Hip Areal BMD (g/cm <sup>2</sup> )	0.82 (0.14)	0.70 (0.11)	< 0.0001
<i>Men</i>			
Sample size	377	63	
Age (years)	75.2 (5.4)	80.2 (5.6)	< 0.0001
BMI (kg/m <sup>2</sup> )	26.8 (3.7)	25.9 (3.9)	0.35
Strength (N)	5140 (1200)	3860 (940)	< 0.0001
Load-to-Strength Ratio	0.82 (0.20)	1.04 (0.24)	< 0.0001
Femoral Neck Areal BMD (g/cm <sup>2</sup> )	0.79 (0.13)	0.65 (0.09)	< 0.0001
Total Hip Areal BMD (g/cm <sup>2</sup> )	0.97 (0.16)	0.81 (0.13)	< 0.0001

<sup>a</sup> p-values comparing cases versus controls are from a T-test for age, and from an age-adjusted F-test for all other comparisons.



**Table 4**

Odds ratio per standard deviation change for prediction of vertebral fracture (and 95% confidence intervals).

	(No Adjustment)		(Adjusted for age)		(Adjusted for age, BMI and prior vertebral fracture)	
	Women	Men	Women	Men	Women	Men
<i>SQ1, SQ2, SQ3 Vertebral Fracture</i>						
Strength	3.1 (2.1-4.7)	2.3 (1.6-3.4)	2.8 (1.8-4.3)	2.2 (1.5-3.2)	2.3 (1.5-3.6)	2.0 (1.3-3.1)
Load-to-Strength Ratio	1.6 (1.3-1.9)	2.0 (1.4-2.7)	1.5 (1.2-1.8)	1.9 (1.4-2.7)	1.4 (1.2-1.8)	1.9 (1.4-2.8)
Vertebral Trabecular BMD	2.4 (1.8-3.2)	1.9 (1.3-2.7)	2.3 (1.7-3.2)	1.7 (1.2-2.5)	1.8 (1.3-2.6)	1.5 (1.0-2.2)
<i>SQ2, SQ3 Vertebral Fracture</i>						
Strength	4.9 (2.8-8.6)	2.5 (1.4-4.4)	4.3 (2.4-7.6)	2.4 (1.3-4.3)	3.3 (1.8-5.9)	1.8 (1.0-3.4) <sup>a</sup>
Load-to-Strength Ratio	1.8 (1.5-2.3)	1.9 (1.2-3.1)	1.7 (1.4-2.2)	1.8 (1.1-3.0)	1.7 (1.3-2.3)	1.6 (0.9-2.8) <sup>a</sup>
Vertebral Trabecular BMD	3.3 (2.2-4.7)	1.5 (0.9-2.5) <sup>a</sup>	3.1 (2.1-4.7)	1.4 (0.9-2.3) <sup>a</sup>	2.5 (1.6-3.7)	1.0 (0.6-1.6) <sup>a</sup>

<sup>a</sup>Not statistically significant (p > 0.05).

**Table 5**

Odds ratio per standard deviation change for prediction of hip fracture (and 95% confidence intervals).

	(No Adjustment)		(Adjusted for age)		(Adjusted for age and BMI)	
	Women	Men	Women	Men	Women	Men
Strength	6.3 (4.0–10)	4.1 (2.8–6.2)	4.2 (2.6–6.9)	3.5 (2.3–5.3)	4.3 (2.6–7.4)	3.7 (2.4–5.7)
Load-to-Strength Ratio	1.9 (1.6–2.4)	2.4 (1.8–3.2)	1.7 (1.4–2.2)	2.3 (1.7–3.1)	2.3 (1.8–3.0)	2.6 (1.9–3.5)
Femoral Neck Areal BMD	3.7 (2.6–5.3)	4.3 (2.9–6.3)	2.7 (1.9–3.9)	3.7 (2.5–5.6)	2.7 (1.8–4.0)	4.0 (2.6–6.1)
Total Hip Areal BMD (g/cm <sup>2</sup> )	3.5 (2.5–5.0)	3.1 (2.2–4.4)	2.6 (1.8–3.7)	2.6 (1.8–3.7)	2.6 (1.8–3.8)	2.8 (1.9–4.1)