ORIGINAL ARTICLE

JBMR®

Microarchitecture and Peripheral BMD are Impaired in Postmenopausal White Women With Fracture Independently of Total Hip *T*-Score: An International Multicenter Study

Stephanie Boutroy,[1,2] Sundeep Khosla,[3] Elisabeth Sornay-Rendu,[2] Maria Belen Zanchetta,[4] Donald J McMahon,[1] Chiyuan A Zhang,[1] Roland D Chapurlat,[2] Jose Zanchetta,[4] Emily M Stein,[1] Cesar Bogado,[4] Sharmila Majumdar,[5] Andrew J Burghardt,[5] and Elizabeth Shane[1]

[1]College of Physicians and Surgeons, Columbia University Medical Center, New York, NY, USA
[2]INSERM UMR1033, Université de Lyon, Hospices Civils de Lyon, Lyon, France
[3]Endocrine Research Unit, College of Medicine, Mayo Clinic, Rochester, MN, USA
[4]Instituto de Diagnóstico e Investigaciones Metabolicas (IDIM), Universidad del Salvador, Buenos Aires, Argentina
[5]Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA, USA

ABSTRACT

Because single-center studies have reported conflicting associations between microarchitecture and fracture prevalence, we included high-resolution peripheral quantitative computed tomography (HR-pQCT) data from five centers worldwide into a large multicenter analysis of postmenopausal women with and without fracture. Volumetric BMD (vBMD) and microarchitecture were assessed at the distal radius and tibia in 1379 white postmenopausal women (age 67 ± 8 years); 470 (34%) had at least one fracture including 349 with a major fragility fracture. Age, height, weight, and total hip *T*-score differed across centers and were employed as covariates in analyses. Women with fracture had higher BMI, were older, and had lower total hip *T*-score, but lumbar spine *T*-score was similar between groups. At the radius, total and trabecular vBMD and cortical thickness were significantly lower in fractured women in three out of five centers, and trabecular number in two centers. Similar results were found at the tibia. When data from five centers were combined, however, women with fracture had significantly lower total, trabecular, and cortical vBMD (2% to 7%), lower trabecular number (4% to 5%), and thinner cortices (5% to 6%) than women without fracture after adjustment for covariates. Results were similar at the radius and tibia. Similar results were observed with analysis restricted to major fragility fracture, vertebral and hip fractures, and peripheral fracture (at the radius). When focusing on osteopenic women, each SD decrease of total and trabecular vBMD was associated with a significantly increased risk of major fragility fracture (OR = 1.55 to 1.88, $p < 0.01$) after adjustment for covariates. Moreover, trabecular architecture modestly improved fracture discrimination beyond peripheral total vBMD. In conclusion, we observed differences by center in the magnitude of fracture/nonfracture differences at both the distal radius and tibia. However, when data were pooled across centers and the sample size increased, we observed significant and consistent deficits in vBMD and microarchitecture independent of total hip *T*-score in all postmenopausal white women with fracture and in the subgroup of osteopenic women, compared to women who never had a fracture. © 2016 American Society for Bone and Mineral Research.

KEY WORDS: HR-PQCT; OSTEOPOROSIS; BONE MICROSTRUCTURE; FRACTURE RISK; MULTICENTER STUDIES

Introduction

Bone microarchitecture is a determinant of bone strength and its measurement may improve fracture risk prediction. Recent advances in imaging technology enable bone microarchitecture to be assessed in vivo, noninvasively, by high-resolution peripheral quantitative computed tomography (HR-pQCT). To date, several single-center cross-sectional studies have reported associations between fracture status and parameters of bone microarchitecture, but those studies have reported conflicting results, with differences observed only at the radius, at both sites, depending on the fracture's site and severity or no difference at all.[1–10]

Received in original form June 12, 2015; revised form December 23, 2015; accepted December 29, 2015. Accepted manuscript online January 28, 2016.
Address correspondence to: Stephanie Boutroy, PhD, INSERM UMR1033, Hôpital Edouard Herriot – Pavillon F, 69437 Lyon Cedex 03, France.
E-mail: stephanie.boutroy@inserm.fr
Additional Supporting Information may be found in the online version of this article.

Journal of Bone and Mineral Research, Vol. 31, No. 6, June 2016, pp 1158–1166
DOI: 10.1002/jbmr.2796

A likely source of these discrepancies is the limited number of fractures in each study (usually fewer than 100 women with prevalent fracture). To overcome the difficulty of enrolling thousands of patients per center, multicenter collaboration is required. So far, most multicenter studies with HR-pQCT focused on interventional/pharmaceutical efficacy with comparisons based on individuals' follow-up changes. Although variability across different devices and imaging centers is important in these settings, it is probably even more crucial when cross-sectional data are compared, because these comparisons are based on absolute values rather than more stable change scores. Single-center reproducibility has been well established, varying from 0.2% to 1.5% for volumetric BMD (vBMD), 0.4% to 5.1% for cortical thickness, and 0.3% to 5.8% for trabecular microstructure.[1,11–15]

Engelke and colleagues[16] have reported short-term in vivo precision data pooled from nine imaging centers. In a context of a multicenter clinical trial, with centralized training and scan analysis, they observed that precision was comparable to single-center results previously reported. Moreover, to fully investigate bone microstructure precision measurement across multiple centers, Burghardt and colleagues[17] have developed a set of anthropomorphic microstructure-realistic imaging phantoms composed of human cadaveric distal radius embedded in resin and mounted on brackets to ensure a reproducible positioning and fixation within the scanner. These phantoms were scanned at nine different HR-pQCT centers and multicenter precision errors were generally less than 5% for volumetric density and microstructure parameters. Although these precision errors were comparable with in vivo single-center precision errors, they differed from ex vivo single-center short-term precision by a factor of two to five.[17]

Our aims were to determine in a large international multicenter cohort the magnitude and direction of bone microarchitectural parameter differences in postmenopausal women with and without a history of fracture, and to establish the feasibility of pooling HR-pQCT data across centers to address mechanistic and clinical issues concerning bone quality.

Subjects and Methods

Study design and population

We have included HR-pQCT data from five academic centers in North America, South America, and Europe into a large multicenter analysis of postmenopausal women with and without prevalent fracture. A total of 1544 postmenopausal women were recruited from 2005 to 2011. Data from 35 black, 55 Hispanic, 57 Asian, and 1397 white women were collected. Our study focused on the 1379 white women who had a valid measurement of the lumbar spine or total hip areal BMD (aBMD) and of the distal radius or tibia microstructure (18 women were excluded). Among the 1379 white postmenopausal women included in the analysis, 470 had at least one prevalent fracture. Each participating center is hereafter arbitrary labeled "center A" through "center E."

All fragility fractures associated with a trauma equivalent to a fall of standing height or less were recorded. Major fragility fractures included fracture of the forearm, humerus, hip, and spine.

Bone densitometry and microarchitecture

aBMD was acquired at the lumbar spine (L_1–L_4) and total hip by DXA (Hologic, Bedford, MA, USA, or GE Lunar, Madison, WI, USA) and expressed as a *T*-score offset from expected peak bone mass as contained in the DXA scanner manufacturers' database. The reference values were country-specific. Using the WHO classification,[18–20] these women were classified as normal (*T*-score ≥ -1), osteopenic ($-1 >$ *T*-score > -2.5), or osteoporotic (*T*-score ≤ -2.5) based on the values of their aBMD measurements at the lumbar spine or total hip.

Volumetric density and bone microarchitecture were assessed at the distal radius and tibia by HR-pQCT (XtremeCT, Scanco Medical AG, Brüttisellen, Switzerland) according to the manufacturer's standard in vivo acquisition protocol.[1,21] The measurement region was manually defined by the operator by placing a reference line at the endplate of the radius and tibia on a preliminary performed anteroposterior scout view. The first CT slice was 9.5 mm and 22.5 mm proximal to the reference line for the distal radius and distal tibia, respectively.

The following imaging settings were used in all centers: effective energy = 60 kVp, X-ray tube current = 900 μA, integration time = 100 ms, 108 projection radiographs on a 180-degree rotation. The 126-mm field of view was reconstructed on a 1536×1536 matrix, yielding 82-μm isotropic voxels.

Methods used to process the CT data have been previously described in detail.[21] Briefly, a trained operator generates semiautomatic contours around the periosteal surface and the entire volume of interest is thereafter automatically separated into a cortical and trabecular region. The outcome variables used in our analyses included total area (Tt.Ar, mm^2, representing the average cross-sectional area of the bone), volumetric bone density (mg hydroxyapatite [HA]/cm^3) for total (Tt.BMD), trabecular (Tb.BMD), and cortical (Ct.BMD) compartments; cortical thickness (Ct.Th, μm); and trabecular number (Tb.N, mm^{-1}), thickness (Tb.Th, μm), separation (Tb.Sp, μm), and intraindividual distribution of separation (Tb.Sp.SD, μm)

Quality assurance procedures for HR-pQCT measurement

In each center, the manufacturer's standard quality protocol was used to monitor device stability. This consisted of daily scans of a phantom containing rods of HA (densities of 0, 100, 200, 400, and 800 mg HA/cm^3) embedded in a soft-tissue equivalent resin (QRM GmbH, Möhrendorf, Germany). As part of the standard quality control procedure, the recommended tolerance was 1% of the highest density rod; ie, within −8 and +8 mg HA/cm^3. Attenuation data were converted to equivalent HA densities.

The quality of HR-pQCT scans was reviewed by an experienced operator in each center and scans with poor quality, as per the manufacturer visual grading of image quality, were excluded from the analyses.

HR-pQCT cross-calibration

A set of four anthropomorphic imaging phantoms, composed of 14 human cadaveric distal radius and five distal tibia, was measured in each center. Analysis of all phantom data was centralized in a single organizing center following the protocol detailed by Burghardt and colleagues.[17] All parameters from the five HR-pQCT scanners were compared for all bone sections. For each parameter, the measured value of each center was normalized to the mean across all centers, and the mean normalized value over all radius or tibia bone sections was calculated.

In each center, patient data were standardized by dividing their density and microstructural parameters by the center corresponding mean normalized value over all radius or tibia.

Statistical analysis

Characteristics are presented as means ± SD and percent difference between group means of women with and without fracture, unless otherwise stipulated, and compared by Student's *t* test or Wilcoxon signed rank test, depending on the distribution of the variables.

One-way analysis of variance by general linear models was used to identify age, height, weight, and total hip *T*-score as between-center differences that would be used as covariates with center interactions. Comparisons between women with and without fracture were performed using logistic regression for all types of fractures and for major fragility fracture. Odds ratios with 95% confidence intervals (ORs [95% CIs]) were computed per SD decrease of density and structural parameters, unless otherwise specified. The Benjamini-Hochberg procedure was used to control for the false discovery rate (FDR).[22] The area under the receiver operating characteristic (ROC) curve (AUC) of volumetric density and structural parameters were compared with AUC of total hip *T*-score after adjustment for covariates. Analyses were independently repeated for peripheral fracture and for vertebral and hip fracture. In addition, we also conducted the analysis with radius (or tibia) total vBMD as covariate (instead of total hip *T*-score) to test whether peripheral trabecular architecture improves fracture discrimination beyond peripheral BMD.

A leave-one-out sensitivity analysis of centers was performed that provided similar data to the one presented (data not shown).

For graphical visualization, the measured difference between fractured and nonfractured subjects of each parameter after adjustment for covariates was plotted by center and for the overall population with 95% CI.

In addition, women were categorized according to the WHO classification as normal, osteopenic, and osteoporotic. In the subgroup of osteopenic women, the associations between fracture and volumetric density and structural parameters were repeated as described above on the whole population.

All statistical analyses were performed using SAS 9.2 (SAS Institute, Cary, NC, USA), SPSS 18.0 software (IBM Corp., Armonk, NY, USA), and/or MedCalc 15.2.2 (MedCalc; https://www.medcalc.org/).

Results

Population characteristics

Among the 1379 white postmenopausal women (age 66 ± 9 years), 470 (34%) had at least one prevalent fracture (194 forearm, 178 vertebral, 35 humerus, 20 hip, 83 lower leg, 46 ribs, 61 metatarsal, 35 metacarpal, 8 clavicle, and 6 patella); 349 had a major fragility fracture, and 170 sustained several fractures.

Each center recruited between 102 and 587 women, with a fracture prevalence ranging from 12% to 75% (Fig. 1).

A total of 292 women: 161 (17.7%) without fracture and 131 (27.9%) with fracture received treatment affecting bone metabolism for 1 year or more at the time of (or within 6 months before) HR-pQCT acquisition ($p < 0.001$; bisphosphonates: 41/77; hormone replacement therapy: 68/31; selective estrogen receptor modulator: 14/11; corticosteroids: 6/3; thyroxine: 17/8; tibolone: 7/1; aromatase inhibitors: 8/0; respectively).

vBMD and microstructural data were missing for 20 radius and 15 tibia because of poor HR-pQCT quality scans.

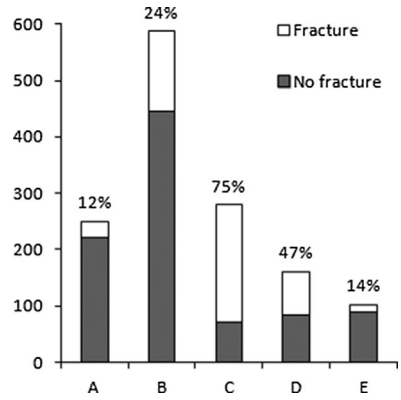

Fig. 1. Sample size by fracture status per center.

Cross-calibration

The variability of volumetric density and structural parameters of phantoms across centers for radius bone sections is shown in Fig. 2 and correction factors per parameter, site and center are summarized in Supporting Table 1.

Across centers, volumetric densities varied between −3.6% and 4.5% of the mean, Ct.Ar and Ct.Th varied within ±2.1% of the mean, and Tb.Ar varied within ±1.3%. Trabecular parameters showed a higher variation across centers between −4.8% and 4.7% at the radius and between −7.5% and 8.2% at the tibia, with the exception of Tb.Sp.SD at the tibia spanning −8.9% and 15.9% of the mean.

Overall and within-center difference between women with and without fracture

Overall and before any adjustment, women with fracture were older by an average 4 years, heavier (4.2%), had a higher BMI (3.1%) and lower total hip *T*-score (−0.3 SD) than women without fracture (all $p \leq 0.01$). Height and lumbar spine *T*-score were not different between fractured and nonfractured women (Table 1). Within centers, these anthropometric differences between fractured and nonfractured women varied as indicated by the range shown in Table 1. Following analyses were therefore performed with age, height, weight, total hip *T*-score, center, and center interactions as covariates.

As illustrated on Fig. 3, at the distal radius, total and trabecular vBMD were significantly lower in fractured women in three out of five centers (by 8% to 11% and 13% to 15%, respectively, $p < 0.05$), whereas total and trabecular area did not differ between fractured and nonfractured women (except in one center for Tb.Ar). Cortical thickness, area, and vBMD were lower in all centers, but the difference reached significance in three of five, respectively (Ct.Th −6% to −12%, Ct.Ar −6% to −11%, $p < 0.05$) and one center for Ct.BMD (−3.5%, $p < 0.05$). Trabecular number and thickness were significantly lower in two centers (by 8% to 10%) and one center (−11%, $p < 0.05$), respectively. Trabecular separation and distribution were significantly higher in two centers (17%, $p < 0.05$) and one center (37%, $p < 0.05$), respectively. Globally, similar results were found at the distal tibia.

When data from five centers were combined, however, women with fracture had significantly lower total, trabecular,

Table 1. Characteristics of the Population

	Women without fracture ($n = 909$)	Women with fracture ($n = 470$)	Difference fracture versus no fracture	Difference range within centers
Age (years)	65 ± 8	69 ± 9	6.4%**	0.7% to 11.1%
Height (cm)	159 ± 7	160 ± 7	0.4%	–0.7% to –0.2%
Weight (kg)	65 ± 12	68 ± 14	4.2%**	–3.3% to 5.6%
BMI (kg/m^2)	25.8 ± 4.7	26.6 ± 5.3	3.1%*	–2.6% to 6.2%
Total hip *T*-score	-0.9 ± 1.0	-1.2 ± 1.0	–0.3 SD**	–0.8 SD to 0.1 SD
Lumbar spine *T*-score	-1.5 ± 1.3	-1.5 ± 1.4	0.0 SD	–0.8 SD to –0.1 SD

*$p \leq 0.01$.
**$p \leq 0.001$.

and cortical vBMD (by 2% to 7%, $p < 0.05$), fewer trabeculae (lower Tb.N by 4% to 5%, $p < 0.05$), and thinner cortices (lower Ct.Th by 6% and Ct.Ar by 5%, $p < 0.05$) than women without fracture. Results were similar at the distal radius and tibia (Fig. 3), except for total and trabecular area differences between fractured and nonfractured women at the radius.

When the analysis was restricted to major fragility fracture ($n = 349$), differences between women with and without fracture were similar to differences observed with all fractures (Fig. 3).

In multivariate logistic regression models including age, height, weight, total hip *T*-score, center, and center interactions, each SD decrease of total, trabecular, and cortical vBMD, trabecular number and thickness, cortical thickness and area, assessed either at the radius or tibia, were associated with a significantly increased likelihood of major fragility fracture. Moreover, each SD increase in trabecular area, separation, and distribution (heterogeneity of the network, at the radius) were also associated with a significantly increased probability of major fragility fracture. The ORs ranged from 1.22 to 1.95 ($p < 0.05$), and AUCs ranged between 0.845 and 0.857 (Table 2). Each SD decrease of total hip and lumbar spine aBMD were associated with a significantly increased likelihood of major fragility fracture: OR 1.96 (95% CI, 1.61 to 2.39) and 1.51 (95% CI, 1.32 to 1.72), respectively (before any adjustment, OR 1.58 [95% CI, 1.38 to 1.81] and 1.03 [95% CI, 0.94 to 1.13]).

At both sites (radius/tibia), AUC for total and trabecular vBMD (0.855/0.857 and 0.856/0.852), cortical area and thickness (0.849/0.848 and 0.850/0.848), and trabecular separation (0.847/0.846), with age, height, weight, total hip *T*-score, center, and center interactions were higher than AUC for total hip *T*-score with age, height, weight, center, and center interactions (0.839 [95% CI, 0.814 to 0.863]), $p < 0.05$.

When the multivariate logistic regression analysis was restricted to vertebral and hip fracture ($n = 188$), differences between women with and without fracture were similar to differences observed with all fracture or major fragility fracture, except for radius trabecular area, cortical vBMD, and trabecular thickness and for trabecular number at the tibia. The ORs ranged from 1.24 to 2.10 ($p < 0.05$ after FDR correction). In contrast, when the analysis was restricted to peripheral fracture ($n = 282$), differences between women with and without fracture were similar to differences observed with vertebral and hip fracture only at the radius (after FDR correction). At the tibia, each SD decrease of total and trabecular vBMD, cortical area and thickness, and trabecular number and thickness were associated

Table 2. Odds Ratio and Area Under the ROC Curve for Major Fragility Fracture by SD Decrease of the Parameter (Otherwise Stipulated), With Age, Height, Weight, Total Hip *T*-Score, Center, and Center Interactions as Covariates

	Radius		Tibia	
	OR (95% CI)	AUC	OR (95% CI)	AUC
Tt.Ar[a]	1.16 (0.96–1.40)	0.844 (0.820–0.868)	1.28 (1.05–1.56)$^{*,+}$	0.846 (0.821–0.870)
Tb.Ar[a]	1.24 (1.03–1.50)$^{*,+}$	0.845 (0.821–0.869)	1.33 (1.09–1.62)$^{**,++}$	0.846 (0.822–0.870)
Ct.Ar	1.46 (1.18–1.80)$^{***,+++}$	0.849 (0.825–0.872)†	1.55 (1.25–1.92)$^{***,+++}$	0.848 (0.824–0.872)†
Tt.BMD	1.75 (1.41–2.18)$^{***,+++}$	0.855 (0.832–0.878)††	1.95 (1.55–2.44)$^{***,+++}$	0.857 (0.833–0.880)†††
Ct.BMD	1.35 (1.09–1.67)$^{**,++}$	0.847 (0.822–0.871)	1.47 (1.19–1.81)$^{***,+++}$	0.847 (0.823–0.871)
Ct.Th	1.50 (1.22–1.85)$^{***,+++}$	0.850 (0.826–0.873)†	1.54 (1.24–1.90)$^{***,+++}$	0.848 (0.824–0.872)†
Tb.BMD	1.74 (1.42–2.12)$^{***,+++}$	0.856 (0.833–0.879)††	1.66 (1.36–2.03)$^{***,+++}$	0.852 (0.829–0.876)††
Tb.N	1.55 (1.30–1.86)$^{***,+++}$	0.852 (0.828–0.875)†	1.24 (1.02–1.49)$^{*,+}$	0.845 (0.821–0.870)
Tb.Th	1.33 (1.11–1.59)$^{**,++}$	0.848 (0.824–0.872)	1.40 (1.18–1.66)$^{***,+++}$	0.847 (0.823–0.872)
Tb.Sp[a]	1.36 (1.15–1.61)$^{***,+++}$	0.847 (0.823–0.871)†	1.22 (1.03–1.45)$^{*,+}$	0.846 (0.822–0.870)†
Tb.Sp.SD[a]	1.27 (1.08–1.48)$^{**,++}$	0.846 (0.821–0.870)	1.23 (0.97–1.31)	0.845 (0.820–0.869)

[a]Data are presented for each SD increase of the parameter.
*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$.
$^{+}p \leq 0.05$; $^{++}p \leq 0.01$; $^{+++}p \leq 0.001$ after FDR correction.
$^{\dagger}p \leq 0.05$, $^{\dagger\dagger}p \leq 0.01$, $^{\dagger\dagger\dagger}p \leq 0.001$ compared to AUC of total hip *T*-score with age, height, weight, center, and center interactions as covariates (0.839; 95% CI, 0.814–0.863).

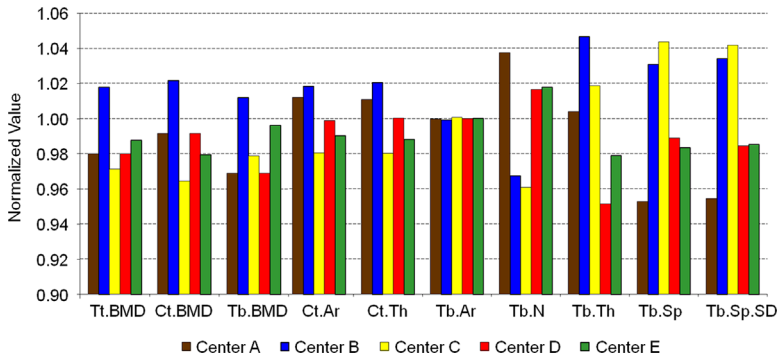

Fig. 2. Mean normalized densitometric and structural value of the radius anthropomorphic imaging phantoms per center.

with a significantly increased likelihood of peripheral fracture. After FDR correction, total and trabecular vBMD, and trabecular thickness remained significantly associated with fracture (ORs ranging from 1.27 to 1.61, $p < 0.03$).

When the multivariate logistic regression analysis was performed with radius (or tibia) total vBMD as covariate (instead of total hip *T*-score), an SD increase in total and trabecular area were associated with a significantly decreased likelihood of major fragility fracture (at the radius: OR 0.69 [95% CI, 0.55 to 0.87] and 0.63 [95% CI, 0.49 to 0.83], $p < 0.01$ after FDR correction; at the tibia: OR 0.72 [95% CI, 0.57 to 0.93] and 0.72 [95% CI, 0.55 to 0.93], $p < 0.05$ after FDR correction) (Table 3). Moreover, at the radius, each SD decrease of trabecular BMD and number were associated with a significantly increased likelihood of major fragility fracture (OR 1.60 [95% CI, 1.25 to 2.06] and 1.42 [95% CI, 1.17 to 1.73], $p < 0.01$ after FDR correction). AUC ranged from 0.852 to 0.858 and did not meaningfully differ from AUC for total vBMD (radius: 0.852 [95% CI, 0.829 to 0.875], tibia: 0.855 [95% CI, 0.832 to 0.878]), when age, height, weight, center, and center interactions were used as covariates.

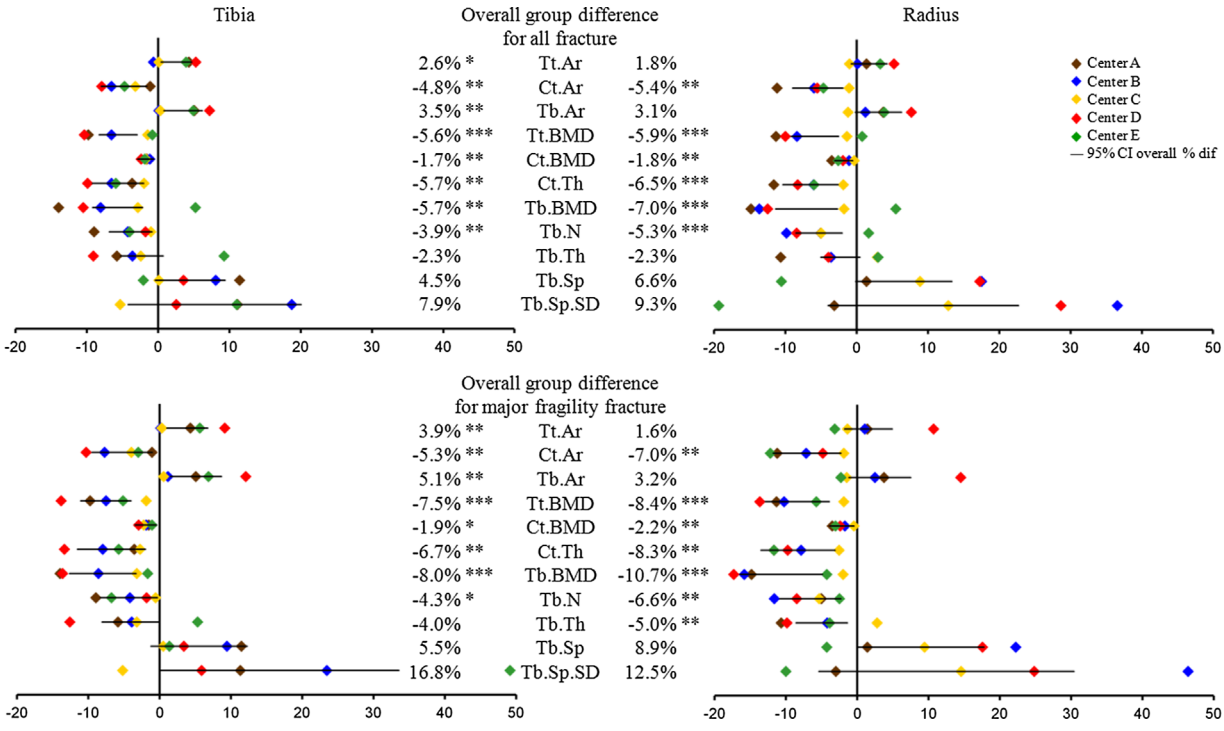

Fig. 3. Overall and within center percentage difference between white women with and without fracture (upper panel) and restricted to major fragility fracture (lower panel). The value of the overall group average percentage difference between the fracture and nonfracture women is shown by number and each line represents the 95% confidence limits surrounding this overall group difference. Each site is placed according to the average percentage difference between fracture and nonfracture women within the site.

Table 3. Odds Ratio and Area Under the ROC Curve for Major Fragility Fracture by SD Decrease of the Parameter (Otherwise Stipulated), With Age, Height, Weight, Total vBMD (Radius or Tibia), Center, and Center Interactions

	Radius		Tibia	
	OR (95% CI)	AUC	OR (95% CI)	AUC
Tt.Ar[a]	0.69 (0.55–0.87)$^{**,++}$	0.856 (0.833–0.878)	0.72 (0.57–0.93)$^{**,+}$	0.858 (0.835–0.881)
Tb.Ar[a]	0.63 (0.49–0.83)$^{***,++}$	0.856 (0.834–0.879)	0.72 (0.55–0.93)$^{*,+}$	0.858 (0.835–0.880)
Ct.Ar	1.00 (0.74–1.34)	0.852(0.829–0.875)	1.14 (0.87–1.48)	0.855 (0.832–0.878)
Ct.BMD	0.80 (0.59–1.08)	0.853 (0.830–0.876)	1.19 (0.94–1.50)	0.856 (0.832–0.879)
Ct.Th	0.78 (0.53–1.15)	0.852 (0.829–0.875)	0.98 (0.72–1.32)	0.855 (0.832–0.878)
Tb.BMD	1.60 (1.25–2.06)$^{***,++}$	0.856 (0.833–0.879)	1.19 (0.90–1.59)	0.856 (0.833–0.879)
Tb.N	1.42 (1.17–1.73)$^{***,++}$	0.854 (0.832–0.877)	0.99 (0.80–1.22)	0.855 (0.832–0.878)
Tb.Th	1.09 (0.89–1.33)	0.853 (0.830–0.876)	1.14 (0.95–1.36)	0.855 (0.832–0.878)
Tb.Sp[a]	1.22 (1.02–1.45)*	0.853 (0.830–0.876)	0.99 (0.82–1.19)	0.855 (0.832–0.878)
Tb.Sp.SD[a]	1.14 (0.97–1.33)	0.853 (0.830–0.875)	0.96 (0.82–1.13)	0.855 (0.832–0.878)

[a]Data are presented for each SD increase of the parameter.
$^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$.
$^{+}p \leq 0.05$; $^{++}p \leq 0.01$; $^{+++}p \leq 0.001$ after FDR correction.

Discrimination of osteopenic women with and without major fragility fracture

Based on the WHO classification, using the lowest of their aBMD measurements at the lumbar spine or total hip, 333 women had normal aBMD (24%), 679 were osteopenic (49%), and 367 were osteoporotic (27%). Among the 349 women with major fragility fracture, 67 were in the normal range of aBMD, 171 were in the osteopenic range, and 111 were in the osteoporotic range with fracture rates respectively of 23%, 28%, and 33%.

Because one-half of the fractures occurred in women in the osteopenic range, analyses of the associations between fracture status and density or structural parameters were repeated in this subgroup. In multivariate logistic regression models including age, height, weight, total hip *T*-score, center, and center interactions, each SD decrease of total and trabecular vBMD assessed at both sites (OR 1.46 to 1.80, $p < 0.03$), was associated with a significantly increased risk of major fragility fracture. In addition, at the radius, each SD decrease of trabecular number (1.37 [95% CI, 1.04 to 1.79]) was associated with a significantly increased risk of major fragility fracture. At the tibia, each SD decrease of cortical thickness (1.40 [95% CI, 1.03 to 1.91]), cortical area (1.44 [95% CI, 1.06 to 1.98]), and cortical vBMD (1.38 [95% CI, 1.01 to 1.87]) were associated with a significantly increased risk of major fragility fracture. After FDR correction, only tibial total vBMD and trabecular vBMD at both sites remained significant.

AUC [95% CI] for tibial total vBMD and trabecular vBMD at both sites (~0.873 [95% CI, 0.843 to 0.903]) with age, height, weight, total hip *T*-score, center, and center interactions, were higher than AUC for total hip *T*-score with age, height, weight, center, and center interactions (0.866 [95% CI, 0.834 to 0.898]).

Discussion

In this large multicenter study, including 470 fractured women among 1379 recruited women, we observed significant and consistent deficits in vBMD and microarchitecture independent of total hip *T*-score at both the distal radius and tibia in postmenopausal white women with fracture compared to women who never had a fracture. These results remained consistent when restricted to major fragility fractures ($n = 349$), vertebral and hip fractures ($n = 188$), and in the subset of women with osteopenia (229 fractures among 679 osteopenic women). In the subset of women with peripheral fracture ($n = 282$), the impairment was mostly observed at the radius. Moreover, peripheral trabecular architecture modestly improved fracture discrimination beyond peripheral BMD (total vBMD measured by HR-pQCT at the radius or tibia).

So far, conflicting results have been reported between fractured and nonfractured women. In a study comparing osteopenic postmenopausal women with ($n = 35$) and without ($n = 78$) prevalent fragility fracture, although spine and hip aBMD were similar, fractured women had lower total and trabecular density and more heterogeneous trabecular distribution ($p < 0.02$) at the radius compared with nonfractured women, but no significant difference at the tibia.[1] In an age-matched case-control study from the OFELY cohort (101 fractured and 101 nonfractured postmenopausal women), vertebral and nonvertebral fragility fractures were associated with low vBMD and architectural deterioration of trabecular and cortical bone at both the radius and tibia, partially independently of aBMD.[3] However, no significant microstructure differences were observed at the distal radius in 36 women with vertebral fracture and 34 controls without osteoporotic fracture from the Rochester Epidemiology project.[7] In another study, postmenopausal women with hip ($n = 62$) or wrist ($n = 50$) fracture had impaired trabecular and cortical compartments at the radius compared to controls ($n = 54$); but only hip fracture subjects had compromised trabecular microstructure at the tibia, and cortical bone was more impaired in hip fracture subjects than in wrist fracture subjects compared to controls.[5] In fractured premenopausal women with idiopathic osteoporosis ($n = 21$), trabecular microstructure was impaired at the radius and tibia but cortical parameters were compromised only at the tibia compared to normal premenopausal women without fracture ($n = 27$).[10] These studies usually enrolled between 20 and 100 fractured women, so that most of them may have been underpowered to provide robust estimates.

In our study, it is reassuring that differences were consistent and of similar magnitude when considering all fractures or only the four major fragility fractures, even if the total number of fractures was reduced by 26%. Moreover, radius and tibia showed a similar pattern, even after adjustment for total hip *T*-score. Based on these observed differences between fractured

and nonfractured women, we were able to calculate the minimal sample sizes to discriminate fracture groups after consideration of age, height, weight, and total hip *T*-score. For an alpha error level of 5% (95% CI) and a statistical power of 80% (beta error level = 20%), the minimal sample sizes per group to discriminate women with major osteoporotic fracture and women without fracture were as follows: 204 and 163 women for total vBMD at the radius and tibia respectively; 202 and 240 women for trabecular vBMD; and 274 and 337 women for cortical thickness. Our study was indeed well powered to observe differences in total and trabecular vBMD as well as cortical thickness, with 470 fractures overall and 349 major fragility fractures. Higher sample sizes were required to discriminate fracture groups with cortical vBMD: 417 and 447 women; trabecular number: 266 and 568, at the radius and tibia, respectively ($\alpha = 5\%$, $\beta = 20\%$).

When the population was restricted to the osteopenic women, which reduced the number of fractures by two-fold, total and trabecular vBMD remained associated with fracture beyond total hip *T*-score. Based on the age-, height-, and weight-adjusted means and SD of our fractured and nonfractured osteopenic women, a sample size of 173 to 325 women per group (depending on the parameter used) would provide robust estimates of the differences observed for radius and tibia total and trabecular vBMD, with an alpha error level of 5% and a statistical power of 80%. With a less conservative statistical power of 50%, while keeping an alpha of 5%, between 76 and 142 women per group would be needed, which is well within our number of osteopenic women with major fragility fracture ($n = 171$) and osteopenic women without fracture ($n = 450$). Discrimination of fracture in this subgroup of osteopenic women is of particular importance because more than one-half of all fractures occur in osteopenic women.[23–25]

When adjusting for radius or tibia total vBMD instead of total hip BMD, only few parameters remained significant: overall geometry (cross-sectional area) and peripheral trabecular architecture (area at the radius and tibia, trabecular vBMD and trabecular number, at the radius). However, because high correlations were reported between total vBMD, cortical thickness, and trabecular parameters (in this study, $r = 0.91$ and 0.81 for Ct.Th at the radius and tibia, and $0.41 \leq |r| \leq 0.76$ for trabecular parameters at both sites), results from Table 3 should be interpreted with caution because of multicollinearity issues. Other statistical approaches (principal component analysis, cluster analysis, etc.) should be investigated to further conclude on the superiority of one parameter to discriminate fracture beyond DXA aBMD or total vBMD measurements. Future research should also address whether the fracture discriminatory capacity of HR-pQCT is different among different categories of patients or if it offers additional discriminatory capacity over aBMD in a particular category of patients.

Bone density and microarchitecture were impaired in fractured women in all centers, but to a variable magnitude. This is likely to be related to the random variability resulting from the limited sample sizes. We cannot rule out, however, that these differences may vary between centers because of selection criteria, lifestyle, environmental, and genetic factors, even if we included only white postmenopausal women to overcome some of these biases. Differences in bone densitometry between countries and continents have been reported and led experts to recommend country specific-normality curves.[26–29] Unfortunately we were not able to cross-calibrate DXA scanners. Moreover, fracture risk varies from one country to another.[30] Difference in height can also have an impact on the measurement region of interest, especially at the distal tibia. Reassuringly, the leave-one-out sensitivity analysis of centers, performed by removing one center at a time, to see the real contribution of each center to the final analysis, provided similar data to the global analysis presented (data not shown).

The comparison of patient data among different HR-pQCT scanners might be controversial because there is no commonly accepted cross-calibration procedure. With our present method —based on the normalization by the mean value of a set of phantoms scanned across all centers—we achieved comparability of HR-pQCT results. Improvement in the cross-calibration universality could be obtained with a better characterization of the phantoms microstructure; ie, not dependent on the mean across all centers but rather from measurements with gold standard μCT systems. Further improvement might be obtained with a better characterization and control of non-device-specific factors known to contribute to measurement variability, such as subject motion and operator biases in patient and reference line positioning.[31–36] The role of the operator is critical for precision but is difficult to evaluate. Establishment of image quality grading scales has helped the operator in his decision making to proceed to the scan analysis or to perform another scan. Pialat and colleagues[35] reported intrareader and interreader agreement on image quality grading with a 5 grade scale, Cohen's kappa score being, respectively, 0.57 for intra-agreement and 0.68 to 0.74 for inter-agreement. The moderately high interreader disagreement highlights the subjective nature of the grading system, even if reassuringly almost all discrepancies were within 1 grade-level.[35] Efforts to develop nonsubjective and quantitative estimates of motion error are clearly needed.[33,36] Moreover, variability of the region of interest positioning following landmark identification on the scout view performed prior to the scan impacts bone measurements.[31,32]

Recently, Bonaretti and colleagues[37] developed an acquisition interface of the HR-pQCT system to simulate the landmark identification by the operator. They evaluated the intraoperator and interoperator variability of the positioning on bone density and microstructure parameters (intraoperator: 7 operators, 15 images positioned 3 times in a random order and interoperator: 5 operators). They observed that positioning was highly reproducible at the tibia in both intraoperator and interoperator configuration, precision (CV_{rms}) was <0.8% and <1.8%, respectively. At the radius, positioning errors were significantly higher, resulting in a precision of bone parameters between 0.4% and 2.1% and 1.0% and 6.6% for intraoperator and interoperator. Highest variability was measured for cortical thickness, with CV_{rms} 1.4-fold to 6.5-fold higher than other parameters. Indeed, Ct.Th has been shown to have the highest variation within the region of interest, with an average thickness three times higher on the 20 most proximal slices than on the 20 most distal, with the standard evaluation method.[2,31,32]

In addition, the image analysis technique may also play a role in variability[13,38,39] with the current standard being a semi-manual contouring. These procedures would further benefit from more rigorous multicenter training procedures and automation of scan positioning and image quality control. Moreover, other microstructural and biomechanical parameters, such as cortical porosity, trabecular topology, stiffness, or failure load that have been reported to be associated with fracture,[6,40–44] but were not evaluated in this study, might also be of interest to discriminate women with and without fracture.

In conclusion, the large sample size available from pooling across centers shows a significant and consistent pattern of

deficits in vBMD and cortical and trabecular microarchitecture that is independent of total hip *T*-score in postmenopausal white women with prevalent fragility fracture. Osteopenic women with and without fracture were also discriminated by total and trabecular vBMD beyond total hip aBMD. We also observed significant heterogeneity in the estimate of these differences across centers. Future research will address the sources of this variability and develop procedural solutions that minimize its influence.

Disclosures

All authors state that they have no conflicts of interest.

Acknowledgments

This research was supported by grants from the NIH (AR052665 to ES; AR058004 to ES; R01-AG17762 to SM; AR027065 to SK; AR060700 to AJB). In addition, this work was supported by an Amgen Fellowship Award and an Institut Servier Postdoctoral Fellowship Award (to SB).

Authors' roles: Study design and conduct: SB, DJM, and ES. Patient recruitment and data interpretation: SB, SK, ESR, MBZ, DJM, ACZ, RDC, JZ, EMS, CB, SM, AJB, and ES. Drafting manuscript: SB, DJM, RC, and ES. Revising manuscript content: SB, SK, ESR, MBZ, DJM, ACZ, RDC, JZ, EMS, CB, SM, AJB, and ES. SB and DJM take responsibility for the integrity of the data analysis.

References

1. Boutroy S, Bouxsein ML, Munoz F, Delmas PD. In vivo assessment of trabecular bone microarchitecture by high-resolution peripheral quantitative computed tomography. J Clin Endocrinol Metab. 2005;90(12):6508–15.
2. Boutroy S, Van Rietbergen B, Sornay-Rendu E, Munoz F, Bouxsein ML, Delmas PD. Finite element analysis based on in vivo HR-pQCT images of the distal radius is associated with wrist fracture in postmenopausal women. J Bone Miner Res. 2008;23(3):392–9.
3. Sornay-Rendu E, Boutroy S, Munoz F, Delmas PD. Alterations of cortical and trabecular architecture are associated with fractures in postmenopausal women, partially independent of decreased BMD measured by DXA: the OFELY study. J Bone Miner Res. 2007;22(3):425–33.
4. Sornay-Rendu E, Cabrera-Bravo JL, Boutroy S, Munoz F, Delmas PD. Severity of vertebral fractures is associated with alterations of cortical architecture in postmenopausal women. J Bone Miner Res. 2009;24(4):737–43.
5. Vico L, Zouch M, Amirouche A, et al. High-resolution pQCT analysis at the distal radius and tibia discriminates patients with recent wrist and femoral neck fractures. J Bone Miner Res. 2008;23(11):1741–50.
6. Melton LJ 3rd, Christen D, Riggs BL, et al. Assessing forearm fracture risk in postmenopausal women. Osteoporosis Int. 2010;21(7):1161–9.
7. Melton LJ 3rd, Riggs BL, Keaveny TM, et al. Structural determinants of vertebral fracture risk. J Bone Miner Res. 2007;22(12):1885–92.
8. Melton LJ 3rd, Riggs BL, Keaveny TM, et al. Relation of vertebral deformities to bone density, structure, and strength. J Bone Miner Res. 2010;25(9):1922–30.
9. Melton LJ 3rd, Riggs BL, van Lenthe GH, et al. Contribution of in vivo structural measurements and load/strength ratios to the determination of forearm fracture risk in postmenopausal women. J Bone Miner Res. 2007;22(9):1442–8.
10. Cohen A, Liu XS, Stein EM, et al. Bone microarchitecture and stiffness in premenopausal women with idiopathic osteoporosis. J Clin Endocrinol Metab. 2009;94(11):4351–60.
11. Burghardt AJ, Buie HR, Laib A, Majumdar S, Boyd SK. Reproducibility of direct quantitative measures of cortical bone microarchitecture of the distal radius and tibia by HR-pQCT. Bone. 2010;47(3):519–28.
12. Kazakia GJ, Hyun B, Burghardt AJ, et al. In vivo determination of bone structure in postmenopausal women: a comparison of HR-pQCT and high-field MR imaging. J Bone Miner Res. 2008;23(4):463–74.
13. MacNeil JA, Boyd SK. Improved reproducibility of high-resolution peripheral quantitative computed tomography for measurement of bone quality. Med Eng Phys. 2008;30(6):792–9.
14. Mueller TL, Stauber M, Kohler T, Eckstein F, Muller R, van Lenthe GH. Non-invasive bone competence analysis by high-resolution pQCT: an in vitro reproducibility study on structural and mechanical properties at the human radius. Bone. 2009;44(2):364–71.
15. Khosla S, Riggs BL, Atkinson EJ, et al. Effects of sex and age on bone microstructure at the ultradistal radius: a population-based noninvasive in vivo assessment. J Bone Miner Res. 2006;21(1):124–31.
16. Engelke K, Stampa B, Timm W, et al. Short-term in vivo precision of BMD and parameters of trabecular architecture at the distal forearm and tibia. Osteoporosis Int. 2012;23(8):2151–8.
17. Burghardt AJ, Pialat JB, Kazakia GJ, et al. Multicenter precision of cortical and trabecular bone quality measures assessed by high-resolution peripheral quantitative computed tomography. J Bone Miner Res. 2013;28(3):524–36.
18. [No authors listed]. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. World Health Organ Tech Rep Ser. 1994;843:1–129.
19. [No authors listed]. Prevention and management of osteoporosis. World Health Organ Tech Rep Ser. 2003;921:1–164.
20. Kanis JA; on behalf of the World Health Organization Scientific Group. Assessment of osteoporosis at the primary health care level. Technical Report. World Health Organization Collaborating Centre for Metabolic Bone Diseases, University of Sheffield, UK; 2007 [cited 2016 Feb 11]. Available from: https://www.shef.ac.uk/FRAX/pdfs/WHO_Technical_Report.pdf.
21. Laib A, Hauselmann HJ, Ruegsegger P. In vivo high resolution 3D-QCT of the human forearm. Technol Health Care. 1998; 6(5–6):329–37.
22. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995;57(1):289–300.
23. Schuit SC, van der Klift M, Weel AE, et al. Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. Bone. 2004;34(1):195–202.
24. Siris ES, Miller PD, Barrett-Connor E, et al. Identification and fracture outcomes of undiagnosed low bone mineral density in postmenopausal women: results from the National Osteoporosis Risk Assessment. JAMA. 2001;286(22):2815–22.
25. Stone KL, Seeley DG, Lui LY, et al. BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. J Bone Miner Res. 2003;18(11):1947–54.
26. Kaptoge S, da Silva JA, Brixen K, et al. Geographical variation in DXA bone mineral density in young European men and women. Results from the Network in Europe on Male Osteoporosis (NEMO) study. Bone. 2008;43(2):332–9.
27. Lunt M, Felsenberg D, Adams J, et al. Population-based geographic variations in DXA bone density in Europe: the EVOS Study. European Vertebral Osteoporosis. Osteoporosis Int. 1997;7(3):175–89.
28. Paggiosi MA, Glueer CC, Roux C, et al. International variation in proximal femur bone mineral density. Osteoporosis Int. 2011;22(2):721–9.
29. Pearson J, Ruegsegger P, Dequeker J, et al. European semi-anthropomorphic phantom for the cross-calibration of peripheral bone densitometers: assessment of precision accuracy and stability. J Bone Miner Res. 1994;27(2):109–20.
30. Elffors I, Allander E, Kanis JA, et al. The variable incidence of hip fracture in southern Europe: the MEDOS Study. Osteoporosis Int. 1994;4(5):253–63.
31. Boyd SK. Site-specific variation of bone micro-architecture in the distal radius and tibia. J Clin Densitom. 2008;11(3):424–30.

32. Mueller TL, van Lenthe GH, Stauber M, Gratzke C, Eckstein F, Muller R. Regional, age and gender differences in architectural measures of bone quality and their correlation to bone mechanical competence in the human radius of an elderly population. Bone. 2009;45(5): 882–91.

33. Pauchard Y, Ayres FJ, Boyd SK. Automated quantification of three-dimensional subject motion to monitor image quality in high-resolution peripheral quantitative computed tomography. Phys Med Biol. 2011;56(20):6523–43.

34. Pauchard Y, Liphardt AM, Macdonald HM, Hanley DA, Boyd SK. Quality control for bone quality parameters affected by subject motion in high-resolution peripheral quantitative computed tomography. Bone. 2012;50(6):1304–10.

35. Pialat JB, Burghardt AJ, Sode M, Link TM, Majumdar S. Visual grading of motion induced image degradation in high resolution peripheral computed tomography: impact of image quality on measures of bone density and micro-architecture. Bone. 2012;50(1):111–8.

36. Sode M, Burghardt AJ, Pialat JB, Link TM, Majumdar S. Quantitative characterization of subject motion in HR-pQCT images of the distal radius and tibia. Bone. 2011;48(6):1291–7.

37. Bonaretti S, Holets M, Derrico NP, et al. Intra- and inter-operator variability in HRpQCT scan positioning. J Bone Miner Res. 2014;29 Suppl 1. [Poster session presented at: Annual Meeting American Society for Bone and Mineral Research (ASBMR); 2014 Sep 12–15; Houston, TX, USA; Presentation Number: MO0289]. Available from: http://www.asbmr.org/education/AbstractDetail?aid=e3a0a66c-cdbf-4917-8be4-eb3dc78143b2.

38. Davis KA, Burghardt AJ, Link TM, Majumdar S. The effects of geometric and threshold definitions on cortical bone metrics assessed by in vivo high-resolution peripheral quantitative computed tomography. Calcif Tissue Int. 2007;81(5):364–71.

39. Valentinitsch A, Patsch JM, Deutschmann J, et al. Automated threshold-independent cortex segmentation by 3D-texture analysis of HR-pQCT scans. Bone. 2012;51(3):480–7.

40. Bala Y, Zebaze R, Ghasem-Zadeh A, et al. Cortical porosity identifies women with osteopenia at increased risk for forearm fractures. J Bone Miner Res. 2014;29(6):1356–62.

41. Pialat JB, Vilayphiou N, Boutroy S, et al. Local topological analysis at the distal radius by HR-pQCT: application to in vivo bone micro-architecture and fracture assessment in the OFELY study. Bone. 2012;51(3):362–8.

42. Vilayphiou N, Boutroy S, Sornay-Rendu E, et al. Finite element analysis performed on radius and tibia HR-pQCT images and fragility fractures at all sites in postmenopausal women. Bone. 2010;46(4): 1030–7.

43. Christen D, Melton LJ 3rd, Zwahlen A, Amin S, Khosla S, Muller R. Improved fracture risk assessment based on nonlinear micro-finite element simulations from HRpQCT images at the distal radius. J Bone Miner Res. 2013;28(12):2601–8.

44. Liu XS, Stein EM, Zhou B, et al. Individual trabecula segmentation (ITS)-based morphological analyses and microfinite element analysis of HR-pQCT images discriminate postmenopausal fragility fractures independent of DXA measurements. J Bone Miner Res. 2012;27(2): 263–72.