

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jmbbm

Research Paper

Experimental validation of DXA-based finite element models for prediction of femoral strength

E. Dall'Ara^{a,b,*}, R. Eastell^{a,b}, M. Viceconti^{b,c}, D. Pahr^d, L. Yang^{a,b}

^aDepartment of Oncology and Metabolism, Mellanby Centre for Bone Research, University of Sheffield, UK

^bINSIGNEO Institute for In Silico Medicine, University of Sheffield, UK

^cDepartment of Mechanical Engineering, University of Sheffield, UK

^dInstitute of Lightweight Design and Structural Biomechanics, Vienna University of Technology, Austria

ARTICLE INFO

Article history:

Received 11 December 2015

Received in revised form

11 May 2016

Accepted 2 June 2016

Available online 10 June 2016

Keywords:

Femur

aBMD

DXA

Finite element

Bone strength

Validation

ABSTRACT

Osteoporotic fractures are a major clinical problem and current diagnostic tools have an accuracy of only 50%. The aim of this study was to validate dual energy X-rays absorptiometry (DXA)-based finite element (FE) models to predict femoral strength in two loading configurations.

Thirty-six pairs of fresh frozen human proximal femora were scanned with DXA and quantitative computed tomography (QCT). For each pair one femur was tested until failure in a one-legged standing configuration (STANCE) and one by replicating the position of the femur in a fall onto the greater trochanter (SIDE). Subject-specific 2D DXA-based linear FE models and 3D QCT-based nonlinear FE models were generated for each specimen and used to predict the measured femoral strength. The outcomes of the models were compared to standard DXA-based areal bone mineral density (aBMD) measurements.

For the STANCE configuration the DXA-based FE models ($R^2=0.74$, $SEE=1473N$) outperformed the best densitometric predictor (Neck_aBMD, $R^2=0.66$, $SEE=1687N$) but not the QCT-based FE models ($R^2=0.80$, $SEE=1314N$). For the SIDE configuration both QCT-based FE models ($R^2=0.85$, $SEE=455N$) and DXA neck aBMD ($R^2=0.80$, $SEE=502N$) outperformed DXA-based FE models ($R^2=0.77$, $SEE=529N$). In both configurations the DXA-based FE model provided a good 1:1 agreement with the experimental data ($CC=0.87$ for SIDE and $CC=0.86$ for STANCE), with proper optimization of the failure criteria.

In conclusion we found that the DXA-based FE models are a good predictor of femoral strength as compared with experimental data *ex vivo*. However, it remains to be investigated whether this novel approach can provide good predictions of the risk of fracture *in vivo*.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

*Corresponding author at: Department of Oncology and Metabolism, Mellanby Centre for Bone Research, University of Sheffield, UK. Tel.: +44 114 2226175.

E-mail address: e.dallara@sheffield.ac.uk (E. Dall'Ara).

1. Introduction

Osteoporotic femoral fractures are a major clinical problem with high morbidity and mortality (Cummings and Melton, 2002; Mnif et al., 2009). Improvement of fracture risk prediction is fundamental in order to improve the patients' quality of life and to reduce the costs associated to the ageing of our society.

Bone mineral density (BMD) measured at the hip is considered the main clinical surrogate of bone strength and, therefore, is used to discriminate osteoporotic patients. Dual energy X-rays absorptiometry (DXA) is most frequently used to evaluate hip areal BMD (aBMD) for its low radiation dose and its low cost (Griffith and Genant, 2008). Even though aBMD can be measured in a number of regions of the hip, this technique allows only 2D measurements without information about a number of important structural parameters such as bone microarchitecture and cortical thickness. Furthermore, although BMD is considered as a surrogate of bone strength, it is difficult to associate it to any mechanical properties. In fact, DXA aBMD was found to predict 29–92% of the variation in experimentally measured femoral bone strength (review of the literature in Dall'Ara et al. (2013b)). Therefore, the research community should focus in finding more reliable and robust way to estimate bone strength. Quantitative computed tomography (QCT) can provide 3D distribution of volumetric BMD (vBMD) that can be measured in different femoral compartments (cortical, trabecular, total) (Engelke et al., 2015; Treece et al., 2015). However, this evaluation cannot be done routinely due to the high radiation dose on the patient.

Subject specific finite element (FE) models can estimate the bone strength in simulated loading scenarios by accounting for geometry and BMD distribution. This method has been recently applied to DXA (Ferdous and Luo, 2015; Naylor et al., 2013; Testi et al., 1999, 2002; Yang et al., 2014), to 2D projections from QCT scans (Den Buijs and Dragomir-Daescu, 2011), and to QCT images (Falcinelli et al., 2014; Keyak et al., 2013; Kopperdahl et al., 2014; Zysset et al., 2015) in order to estimate the risk of fracture or the effect of anti-osteoporotic drug treatments. Validation studies that compare the FE outcomes to bone strength measurements performed on cadaveric femora have shown that QCT based FE can predict 80–90% of femoral strength in simulated fall (Dragomir-Daescu et al., 2011; Keyak et al., 1998; Koivumaki et al., 2012; Luisier et al., 2014; Schileo et al., 2014; Zysset et al., 2013) and 80–94% of femoral strength in simulated one legged stance (Schileo et al., 2014; Cody et al., 1999; Dall'Ara et al., 2013a; Hambli and Allaoui, 2013). To the authors' knowledge, there is only one study that compared the DXA-based FE models predictions of bone strength with experimental measurements (Naylor et al., 2013), which showed that the models can predict 59% of variability of femoral strength in a simulated fall configuration. In another study, Den Buijs and Dragomir-Daescu (2011) used projected images from QCT scans of the proximal femur and generated 2D FE models from those images to predict the femoral stiffness in a simulated fall configuration. In that study the stiffness predicted by the FE models explained between 77%

(for a set of samples used to train the model, $N=9$) and 69% (for a validation set of samples, $N=13$) of the measured bone strength. However, the relationship between outputs of 2D FE models generated from projected QCT scans or from DXA images is still unknown. Considering that DXA measurements are performed routinely for assessment of aBMD and discrimination of osteopenic and osteoporotic patients, the DXA based FE method has potential for improving the prediction of risk of fracture, however, at the current stage its moderate relationship with femoral strength is a limitation and need further investigation.

The aim of this study was to evaluate the ability of DXA based 2D FE models in predicting femoral failure load tested in two loading configurations. The results were compared to standard densitometric parameters and to QCT based 3D FE models.

2. Materials and methods

The sample preparation, QCT scanning and mechanical testing have been already described in details in other studies (Dall'Ara et al., 2013a, 2013b; Luisier et al., 2014; Zysset et al., 2013). Even though several different BMD–elasticity relationships are reported in the literature (Helgason et al., 2008; Zysset, 2003), the constitutive models used in the QCT-based and DXA-based FE analyses were the same used in previous publications (Naylor et al., 2013; Dall'Ara et al., 2013a) for easy comparison of the results with the literature. The main steps are briefly reported in the following paragraphs and in Fig. 1.

2.1. Sample preparation

The ethics commission of the Medical University of Vienna approved all procedures applied during the present study (EK Nr 175/2011). Thirty-six pairs of human femora (17 males, 19 females with age 76 ± 12 years, range 46–96) were dissected and kept frozen at -20°C when not handled. Of each pair, one femur was prepared for being tested in a “STANCE” configuration and the contra-lateral for being tested in a “SIDE” configuration. The specimen were randomly assigned to the STANCE or SIDE group with the only constrain to have to have the same number of left or right femora tested in each loading configuration.

2.2. QCT and DXA scanning

Each sample was submerged in 0.9% NaCl saline solution, exposed to vacuum for 10 min to remove air bubbles and scanned with QCT (Brilliance64, Philips, Germany) and DXA (Discovery QDR, Hologic Inc., USA). QCT scans (intensity: 100 mA; voltage: 120 kV; filter type: B (+0.5 enhancement); voxel size: $0.33 \times 0.33 \times 1.0 \text{ mm}^3$) were performed together with a calibration Phantom (BDC Phantom, QMR GmbH, Germany) in order to convert the HU to equivalent BMD scale (in mgHA/cc). aBMD of the total proximal femur (Tot_aBMD), of the femoral neck (Neck_aBMD), of the trochanteric region (Troch_aBMD) and of the intertrochanteric regions (Inter_aBMD) were computed from each DXA scan

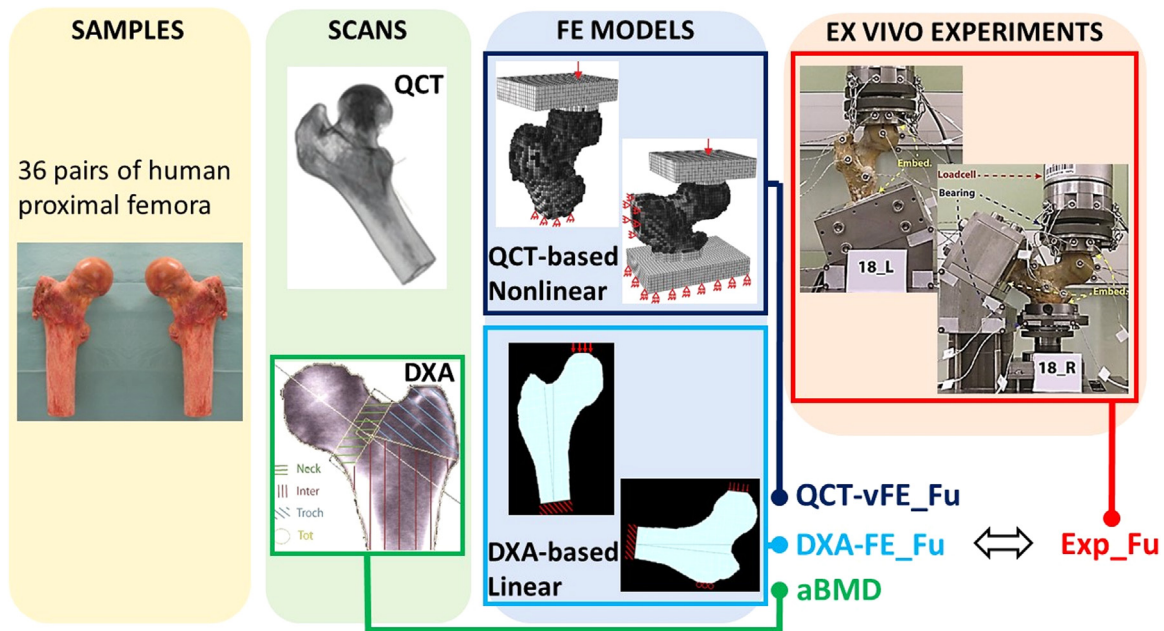


Fig. 1 – Overview of the methods used in this study. Thirty-six pairs of human femora were collected. Each proximal femur was scanned with QCT and DXA. From each pair one sample was tested simulating a one legged stance (STANCE) and the other a sideways fall (SIDE). 3D and 2D FE models were generated from QCT and DXA images, respectively by simulating either a STANCE or a SIDE configuration. The ability of predicting femoral strength of DXA based FE models, QCT based FE models and densitometric parameters was compared for both loading configurations.

(pixel size $0.901 \times 1.000 \text{ mm}^2$) (Dall'Ara et al., 2013b).

2.3. Mechanical tests

For both loading configurations the load was applied in the plane containing both neck and proximal shaft axes and inclined of 60° or 20° from the proximal shaft axis for the SIDE and STANCE configurations, respectively. The shaft was fixed 10 mm below the lesser trochanter and the load was distributed to the superior (STANCE) or medial (SIDE) side of the femoral head with sample specific Polyurethane (PU) caps. For the SIDE configuration, ten millimeters of the lateral part of the greater trochanter were embedded in PU for a proper distribution of the reaction force. The rotation and two translations in the plane perpendicular to the loading and reaction force axis were left free. Tests were performed on a servo-hydraulic testing machine (Mini-Bionix, MTS system, U.S. A.) at a rate of 5 mm/min until failure. The femoral failure load (Fu) was defined as the maximum absolute load along the vertical direction.

2.4. 3D QCT based finite element analysis

The QCT images were resampled to $3 \times 3 \times 3 \text{ mm}^3$ in size voxels. After proper contouring of the proximal femur (Pahr and Zysset, 2009) voxels were directly converted to linear hexahedron elements. PU embedding material (1.36 GPa) and steel loading plate (210 GPa) were modeled. The degrees of freedom allowed in the experiments by the bearings and by the relative rotations allowed between the femoral head cartilage and the subject specific PU embedding were replicated in the models. Bone was considered as heterogeneous and isotropic with an elastic-

damage constitutive law adapted from Garcia et al. (2009) as explained in Dall'Ara et al. (2013a). Due to lack of information about trabecular orientation from the clinical images, the orthotropic model defined in Garcia et al. (2009) was converted to an isotropic model. The average BMD value in each element was converted into bone volume fraction value (Dall'Ara et al., 2013a) that was used to define the material properties of the element. In order to account for nonlinear material behavior, a generalized piecewise Hill yield surface was defined and a scalar damage (in the range 0–1), representing the reduction of the elastic modulus, was modeled. Material properties (elastic and strength) were adjusted from Rincon-Kohli and Zysset (2009), who performed multi-axial mechanical testing on human trabecular bone. To take into account for cortical bone a nonlinear scalar tissue function for both compression and tension was defined to provide an elastic modulus equal to 24 GPa, a compression ultimate stress equal to 266 MPa and a tension ultimate stress equal to 200 MPa for bone without porosities (Bayraktar et al., 2004; Ohman et al., 2011). Analyses were performed (Abaqus 6.11, Simulia, Dassault Systemes, Velizy-Villacoublay, France) until the load–displacement curve reached a clear maximum (vFE_Fu).

2.5. 2D DXA based finite element analysis

The pixel-by-pixel aBMD map was extracted from each DXA scan and each femoral pixel was converted into 2D 4-nodes plane stress elements (element size $0.405 \text{ mm} \times 0.405 \text{ mm}$) by using the procedure reported in Naylor et al. (2013). The material properties in each element were assigned based on a number of assumptions. First, each femur was assumed to be a plate with a subject-specific constant thickness t :

$$t = 3.5\pi W/16$$

where W is the mean width of the middle third cross sections of the femoral neck on the BMD map. The equation imposes that the cross section area and the moments of inertia are as close as possible between the plate's rectangular and the anatomical cross sections, that was assumed circular. Bone was considered as heterogeneous based on aBMD values and isotropic. Poisson ratio was set to 0.3 and Young's modulus and yield stresses in compression and tension were calculated in each element according to the following method. The voxel aBMD (ρ_a) was converted to volumetric BMD (ρ_v) by dividing it by the thickness (t). Local apparent density (ρ_{app}) was computed according to the relationships reported by [Schileo et al. \(2008\)](#).

The mechanical properties in each element were then calculated from the empirical equations provided by [Morgan et al. \(2003\)](#), [Morgan and Keaveny \(2001\)](#):

Modulus of Elasticity (MPa)

$$= \begin{cases} 15010\rho_{app}^{2.18} & \text{if } \rho_{app} \leq 0.280 \text{ g/cm}^3 \\ 6850\rho_{app}^{1.49} & \text{if } \rho_{app} > 0.280 \text{ g/cm}^3 \end{cases}$$

Compressive yield stress S_c (MPa)

$$= \begin{cases} 85.5\rho_{app}^{2.26} & \text{if } \rho_{app} \leq 0.355 \text{ g/cm}^3 \\ 38.5\rho_{app}^{1.48} & \text{if } \rho_{app} > 0.355 \text{ g/cm}^3 \end{cases}$$

Tensile yield stress S_t (MPa)

$$= \begin{cases} 50.1\rho_{app}^{2.04} & \text{if } \rho_{app} \leq 0.355 \text{ g/cm}^3 \\ 22.6\rho_{app}^{1.26} & \text{if } \rho_{app} > 0.355 \text{ g/cm}^3 \end{cases}$$

Tensile yield strain was considered equal to 7300 micro-strain, and compressive yield strain was considered as 10,400 microstrain ([Falcinelli et al., 2014](#); [Bayraktar et al., 2004](#)). To account for the side-artefact errors in biomechanical testing of cadaveric trabecular specimen due to the isolation from its original structure, the above material properties were increased by a factor of 1.28 ([Orwoll et al., 2009](#)).

The loading conditions of the mechanical testing were simulated by distributing an applied load of 2000 N on the flat surface of the modeled cement padding on the femoral head and by constraining the appropriate degrees of freedom on the flat surfaces of the greater trochanter and on the distal end of the femoral shaft.

Failure load (DXA-FE_Fu) from the FE linear analysis was defined as the applied load at which the element stress (or strain) exceeded the failure stress (or strain) for a pre-determined number of elements. For each element within the anatomical region bounded proximally by the subcapital line and distally by a transverse line passing through the distal end of the lesser trochanter, the ratio of the applied stress (or strain) to the yield criterion (defined in the stress or strain space) was computed. Afterwards a contiguous area A within which the ratios were highest was identified as the region where the fracture initiated. Being a linear analyses, the failure force or the femoral strength was calculated by dividing the applied force by the minimum ratio in that area ([Keyak et al., 1998](#)). In order to find the best combination of failure criterion and dimension of the area, we performed an

optimization analyses based on the experimentally measured failure load F_u . We considered different dimensions of A : one element ("global"), one squared millimeters, four square millimeters, nine square millimeters, sixteen square millimeters and 25 square millimeters. Different failure criteria were considered in order to compute the ratio: von Mises stress (VMS), von Mises strain (VME), principal tensile stress (PST), principal compressive stress (PSC), minimum between PST and PSC (PSmin), principal tensile strain (PET), principal compressive strain (PEC), and minimum between the PET and PEC (PEmin).

A suite of Matlab functions, accessible from a Matlab graphic user interface DXA_HipFE, were developed to segment DXA images, generate and solve the FE models and post-process the FE results ([Fig. 2](#)).

2.6. Statistics

For the sensitivity analyses, in order to account both for correlation and quality of the prediction, the concordance correlation coefficient (CC; [Lin, 1989](#)) was computed for the predictions between the DXA-FE_Fu estimated by the different DXA based models and the Exp_Fu. Linear regressions were used to predict the Exp_Fu with DXA-Tot_aBMD, DXA-Neck_aBMD, DXA-Troch_aBMD, DXA-Inter_aBMD, DXA-FE_Fu and QCT-vFE_Fu. Coefficients of determination (R^2) and standard errors of the estimate (SEE) were computed for all predictors. The Cook's distance method was used to study potential outliers in the linear regressions. In particular, we used a mild criterion and removed only those points with Cook's distance larger than ten times the mean Cook's distance value for that regression ([Stevens, 1984](#)).

3. Results

Due to technical problems during the DXA measurements, one pair of specimens had to be removed and the DXA densitometric properties and the DXA based FE models were analyzed for the remaining 35 pairs (16 males, 19 females with age 76 ± 12 years, range 46–96).

The DXA-based FE models matched best the experimental results if the PEC or PE_min with a failure area of 9 mm² were used for the FALL configuration (CC=0.87) or the PEC with failure area of 4 mm² were used for the STANCE configuration (CC=0.86) ([Table 1](#)). Therefore, the PEC criterion with A equal to 9 mm² for SIDE and equal to 4 mm² for STANCE were used for comparing the DXA-FE models with the other predictors.

The best densitometric predictor was Neck_aBMD for both SIDE ($R^2=0.80$ vs $R^2=0.69-0.75$ for the other regions) and STANCE ($R^2=0.66$ vs $R^2=0.54-0.60$ for the other regions) configurations. For SIDE configuration best predictor was found to be QCT-based FE models ($R^2=0.85$, SEE=455N), followed by DXA-Neck_aBMD ($R^2=0.80$, SEE=502N) and 2D DXA-based FE models ($R^2=0.77$, SEE=529N). For STANCE configuration the best predictor was found to be QCT-based FE models ($R^2=0.80$, SEE=1314N), followed by DXA-based FE models ($R^2=0.74$, SEE=1473N) and DXA-Neck_aBMD ($R^2=0.66$, SEE=1687N). Relationships between experimental

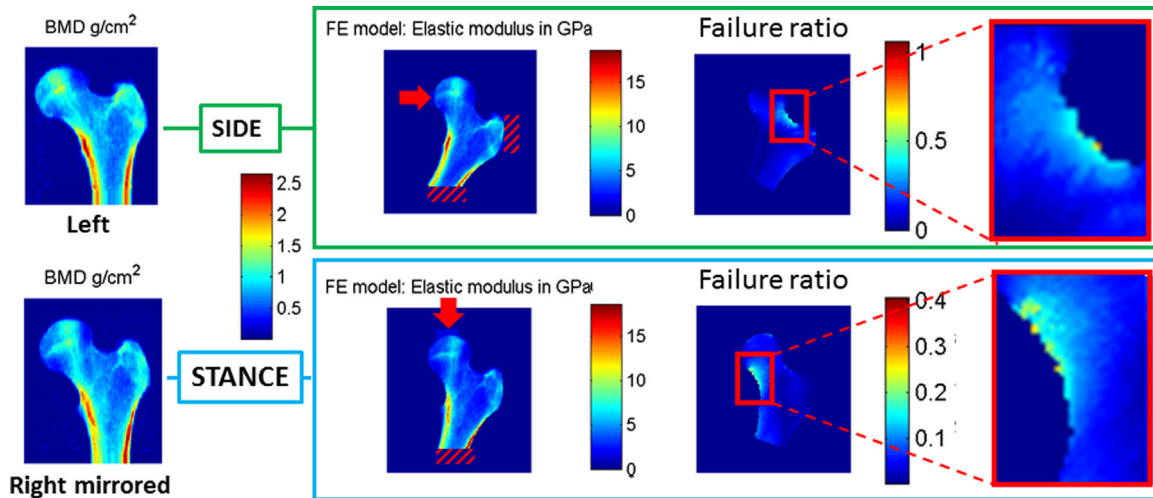


Fig. 2 – Example of the application of DXA-based FE models. From left to right: DXA images of the left and right femur are taken (the image of the right femur was mirrored in order to have a similar orientation), the aBMD distribution is extracted from each image, the image is rotated for assigning the proper boundary conditions for simulated SIDE (top) and STANCE (bottom), the aBMD is converted into elastic modulus distribution, and the failure ratio is computed for each model (in this case a PEC criterion was used).

Table 1 – Concordance coefficients calculated for the prediction of experimental femoral strength from DXA-based FE models generated with different failure criteria: von Mises stress (VMS), von Mises strain (VME), principal tensile stress (PST), principal compressive stress (PSC), minimum between PST and PSC (PSmin), principal tensile strain (PET), principal compressive strain (PEC), and minimum between the PET and PEC (PEmin). The results are reported for different areas where the minimum threshold was reached: single element (global), 1 mm², 2 mm², 9 mm², 16 mm² and 25 mm². The results are reported for both SIDE and STANCE configurations. The highest values for both configurations are reported in bold.

	VMS	VME	PST	PSC	PET	PEC	PS_min	PE_min
SIDE								
Global	0.25	0.15	0.60	0.32	0.58	0.37	0.32	0.37
1 mm ²	0.45	0.28	0.72	0.56	0.86	0.72	0.56	0.72
4 mm ²	0.58	0.39	0.69	0.71	0.82	0.85	0.70	0.85
9 mm ²	0.69	0.49	0.66	0.80	0.72	0.87	0.79	0.87
16 mm ²	0.77	0.55	0.59	0.84	0.63	0.84	0.84	0.84
25 mm ²	0.82	0.63	0.52	0.83	0.54	0.80	0.85	0.80
STANCE								
Global	0.24	0.21	0.42	0.34	0.47	0.59	0.33	0.44
1 mm ²	0.32	0.32	0.60	0.48	0.75	0.80	0.48	0.70
4 mm ²	0.36	0.40	0.57	0.58	0.77	0.86	0.58	0.82
9 mm ²	0.40	0.50	0.47	0.63	0.72	0.84	0.63	0.85
16 mm ²	0.42	0.57	0.36	0.68	0.62	0.79	0.68	0.82
25 mm ²	0.47	0.66	0.31	0.73	0.48	0.72	0.73	0.75

Fu and best aBMD and model predictors are reported in Fig. 3.

Good correlations were found between DXA-based and QCT-based FE models for both STANCE and FALL configurations (for both R²=0.76) as well as between DXA-based FE models and DXA-Neck_aBMD (R²=0.77 for FALL and R²=0.66 for STANCE configurations, Fig. 4).

On similar PC Desktop machines, it typically took less than 2 min to segment a DXA hip scan with reasonable quality and to perform a FE simulation of STANCE or SIDE, whereas the time needed to semi-automatically segment, calibrate, run and post-process the QCT-based FE models was approximately 60 min (approximately 20 min of running time). The typical number of degrees of freedom was 50,000 for DXA-based model and 40,000 for QCT-based models.

4. Discussion

The goal of this study was to investigate the ability of DXA-based FE models in predicting experimentally measured femoral strength in two loading configurations as compared with densitometric parameters and QCT-based FE models.

The results are part of a much larger project where an unique database of experimental and computational results has been created on the same set of 36 pairs of human femora (Dall'Ara et al., 2013a, 2013b; Luisier et al., 2014). In this study we added the estimation of failure load from DXA-based FE models and we compared their output to already published predictions from densitometric measurements and

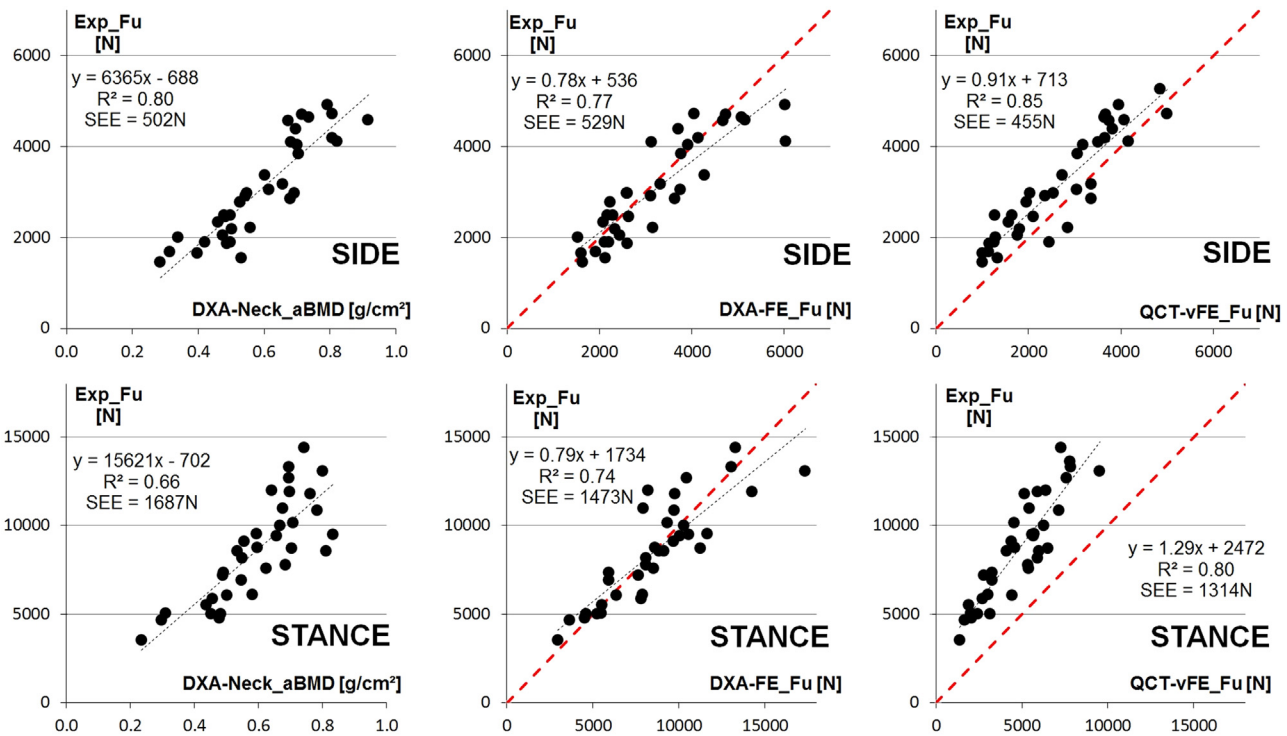


Fig. 3 – Linear regression for the three predictors of experimental femoral strength Exp_Fu for both SIDE (top) and STANCE (bottom) configurations: DXA neck aBMD (left), DXA-based FE model estimations of femoral strength DXA-FE_Fu (center) and QCT-based model estimations of femoral strength QCT-vFE_Fu (right). For each correlation, the regression equation, the coefficient of determination (R^2) and the standard error of the estimate (SEE) are reported.

more complex 3D QCT-based FE models. The strength of the new models is that they are generated from the standard clinical imaging technique (DXA) used routinely for discrimination of osteopenic and osteoporotic patients on the bases of their aBMD. On the other side, these 2D DXA-based linear FE models need a number of simplifications derived from the usage of 2D input images. For example the geometry of the bone is simplified and a constant thickness of the femur is defined based on the width of the femoral neck; the assignment of the material properties is based on the pixel aBMD value which accounts for both cortical and trabecular bone; the boundary conditions can be assigned only in the 2D plane of the DXA image. Interestingly we found that for both SIDE and STANCE configurations the failure criterion that best matched the experiments was based on the minimum compressive strain (PEC). However, the minimum volume of the failed area that best estimated the femoral load was different for the two loading conditions (9 mm² for SIDE and 4 mm² for STANCE). This difference may be due to overloaded regions in the two loading condition, i.e. the inferior side of the femoral neck in STANCE and the superior side of the femoral neck in SIDE configuration.

The DXA-based FE predicted up to 77% and 74% of the variability of the experimental femoral failure load for SIDE and STANCE loading configurations, respectively. The CC equal to 0.87 (SIDE) or 0.86 (STANCE) underline good 1:1 relationships between predicted and experimental measurements if the failure load criterion is optimized for the two loading configurations. The predictions were found to be better compared to the indirect validation performed in the

study of [Naylor et al. \(2013\)](#) for simulated sideways fall configurations (R^2 equal to 0.77, slope equal to 0.78 and intercept equal to 536N in the present study vs R^2 equal to 0.59, slope equal to 0.56 and intercept equal to 1475N in Naylor's study). This improvement is probably due to the better match between the computational and experimental boundary conditions as well as the new optimization of the failure criterion. In particular, the proximal femora were loaded in the plane containing both proximal and femoral neck axis, which was also similar to the plane perpendicular to the DXA scanning. Therefore, the out of axis loads, not reproducible in a 2D model, were minimized in this case. The DXA-based FE models for the SIDE configuration were found to provide similar coefficient of determination when compared to the results reported by [Den Buijs and Dragomir-Daescu \(2011\)](#), who used projected images from QCT scans to estimate the bone stiffness in a simulated fall. While this result may suggest similarities between the predictive ability of models generated from real DXA images or from projected QCT scans, the differences between the models, the studied sample size, and the comparison between predicted and measure mechanical properties recommend further analyses to be done.

For both STANCE and SIDE configurations DXA-based FE models showed lower prediction ability for experimental failure load ($R^2=0.74$ and $R^2=0.77$) worse than the ones of QCT based FE models ($R^2=0.80$ and $R^2=0.85$). This is not surprising considering the much more complex and time consuming 3D models, that include a whole 3D geometry of the proximal femur and the material nonlinearities based on

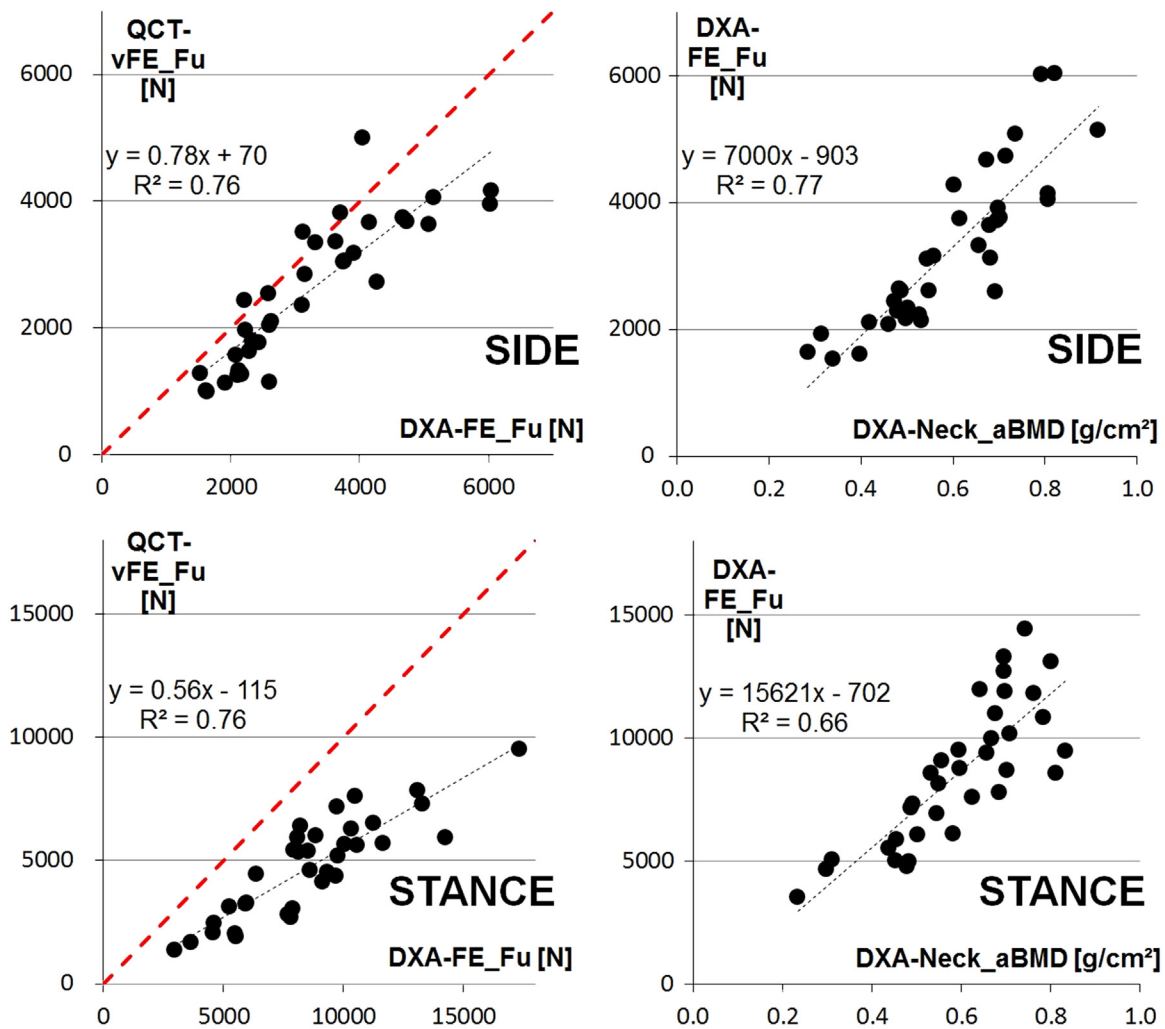


Fig. 4 – Correlations between DXA-based FE model estimations of femoral strength DXA-FE_Fu and DXA neck aBMD (right) and QCT-based model estimations of femoral strength QCT-vFE_Fu (right) for both SIDE (top) and STANCE (bottom) configurations.

the voxel specific calibrated BMD. Predictions of bone strength with the FE models and aBMD were systematically better for the SIDE configuration. This may be due to the more important role that trabecular microarchitecture, not reliably measurable in QCT (Larsson et al., 2014) or DXA images, may have in STANCE loading condition. In fact, the work of Luisier et al. (2014) showed similar predictions of high resolution peripheral QCT (HRpQCT)-based FE models, that include local fabric information, for the STANCE and SIDE configurations in the same set of samples tested in the present study. The DXA based FE were able to predict only 6–8% less of the variability in femoral strength for both configurations, making this tool attractive for further developments. These differences were also underlined by the good correlations between the DXA-based and QCT-based FE models predictions ($R^2=0.76$ for both configurations). The different slopes and intercepts for the two loading configurations are probably due to the fact that the 3D QCT-based FE models were not optimized for the failure criterion but were based on a constitutive model whose material parameters were assessed from experiments on trabecular bone samples and adapted for cortical bone (Dall'Ara et al., 2013a; Rincon-Kohli and Zysset, 2009). It has

to be considered that 3D QCT based FE models have a larger potential for prediction of femoral strength and the simulation of different boundary conditions for estimation of the minimum physiological strength and minimum pathological strength for further predictions of risk of fracture (Falcinelli et al., 2014). Recently a novel approach to generate 3D models from combination of statistical shape models and 2D DXA scans has been developed (Sarkalkan et al., 2014; Vaananen et al., 2015) and could be an appealing solution for the further developments of subject specific models based on clinical DXA images but with 3D applications.

The DXA-based FE models provided better predictions of femoral strength compared to the best densitometric predictor DXA-Neck aBMD only in case of STANCE ($R^2=0.74$ vs $R^2=0.66$), while for SIDE configuration DXA aBMD provided better correlation ($R^2=0.80$ vs $R^2=0.77$). While the lower predictive ability of DXA-based FE compared to DXA-Neck aBMD for SIDE configuration is surprising, it should be noted that two specimens (one for SIDE and one for STANCE) could be considered as outliers (Cook's distance larger than ten times the average value of Cook's distance (Stevens, 1984)) and if excluded the prediction of DXA-FE models improved

(SIDE: $R^2=0.81$, $SEE=488N$, $slope=0.86$, $intercept=328N$, $CC=0.90$; STANCE: $R^2=0.75$, $SEE=1413N$, $slope=0.88$, $intercept=1077N$, $CC=0.86$). The outliers may be related to slightly different positioning of the sample during the DXA scanning that could largely affect the match between experimental and model boundary conditions. With similar criterion no outliers were found for the predictions of QCT-FE F_u and DXA-Neck aBMD. However, for completeness the results were reported for the whole dataset.

The main limitation of the developed DXA-FE models is that the failure load is predicted by using linear models on the bases of a strain criterion, ignoring the nonlinear behavior of the bone before failure. While the tests performed on this dataset show nonlinear behavior especially for the SIDE configuration (Dall'Ara et al., 2013b), this could be due to the low strain rate used during the experiments. In fact, in another study fresh frozen femora tested in physiological conditions and strain rate exhibited a linear behavior until failure (Juszczczyk et al., 2011). Moreover, further investigations are needed in order to provide a failure criterion based on the aBMD values obtained from DXA. Moreover, the DXA-based models were tuned by using the whole set of experiments. While this decision was taken in order to increase the sample size and the range of the variable to be predicted, further tests should be done in order to check if the chosen failure criterion would indeed provide similar results for an independent set of experiments. Also, an optimization of the material parameters for constitutive relationship of the QCT-based models may be necessary. However, this optimization is not trivial due to the complexity of the 3D nonlinear heterogeneous models and future work needs to be done in this topic. Finally, the DXA-based and QCT-based FE models were based on different density-elasticity and density-yield relationships. While this choice was taken in order to be consistent with previously reported procedures, more studies should be performed to investigate how the models would behave if based on the same constitutive law and which one is the optimal for the two approaches.

This is the first study where DXA- and QCT-based FE models were compared for prediction of femoral strength on the same set of experimental data for two different loading configurations. In conclusion the DXA-based FE models were found to predict up to 77% of the femoral strength with a good 1:1 accuracy. Considering the clinical applicability of this method in combination with aBMD, the next step would be to test its potential in improving the ability of discriminating for patients at high or low risk of fracture compared to aBMD alone and for studying the effect of antiresorptive or anabolic drug treatments onto the femoral strength.

Acknowledgments

The authors gratefully acknowledge Prof. Viceconti for the fruitful discussions, Prof. Pretterklieber for providing the tissue, Prof. Kainberger for access to the scanning machines, Prof. Zysset for the nonlinear constitutive law used in the QCT-based models, Mr. Luisier for help during the sample preparation and mechanical testing, and Ms. R. Posch for the

help with the DXA data transfer. The study was partially funded by the EPSRC Frontier Grant (MultiSim project code EP/K03877X/1) and the Arthritis Research UK (code 19669).

REFERENCES

- Bayraktar, H.H., Morgan, E.F., Niebur, G.L., Morris, G.E., Wong, E.K., Keaveny, T.M., 2004. Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. *J. Biomech.* 37, 27–35.
- Cody, D.D., Gross, G.J., Hou, F.J., Spencer, H.J., Goldstein, S.A., Fyhrie, D.P., 1999. Femoral strength is better predicted by finite element models than QCT and DXA. *J. Biomech.* 32, 1013–1020.
- Cummings, S.R., Melton, L.J., 2002. Epidemiology and outcomes of osteoporotic fractures. *Lancet* 359, 1761–1767.
- Dall'Ara, E., Luisier, B., Schmidt, R., Kainberger, F., Zysset, P., Pahr, D., 2013. A nonlinear QCT-based finite element model validation study for the human femur tested in two configurations in vitro. *Bone* 52, 27–38.
- Dall'Ara, E., Luisier, B., Schmidt, R., Pretterklieber, M., Kainberger, F., Zysset, P., Pahr, D., 2013. DXA predictions of human femoral mechanical properties depend on the load configuration. *Med. Eng. Phys.* 35, 1564–1572 discussion 1564.
- Den Buijs, J., Dragomir-Daescu, D., 2011. Validated finite element models of the proximal femur using two-dimensional projected geometry and bone density. *Comput. Methods Prog. Biomed.* 104, 168–174.
- Dragomir-Daescu, D., Op Den Buijs, J., McEligot, S., Dai, Y., Entwistle, R.C., Salas, C., Melton 3rd, L.J., Bennet, K.E., Khosla, S., Amin, S., 2011. Robust QCT/FEA models of proximal femur stiffness and fracture load during a sideways fall on the hip. *Ann. Biomed. Eng.* 39, 742–755.
- Engelke, K., Lang, T., Khosla, S., Qin, L., Zysset, P., Leslie, W.D., Shepherd, J.A., Schousboe, J.T., 2015. Clinical use of quantitative computed tomography (QCT) of the hip in the management of osteoporosis in adults: the 2015 ISCD official positions – Part I. *J. Clin. Densitom.* 18, 338–358.
- Falcinelli, C., Schileo, E., Balistreri, L., Baruffaldi, F., Bordini, B., Viceconti, M., Albisinni, U., Ceccarelli, F., Milandri, L., Toni, A., Taddei, F., 2014. Multiple loading conditions analysis can improve the association between finite element bone strength estimates and proximal femur fractures: a preliminary study in elderly women. *Bone* 67, 71–80.
- Ferdous, Z., Luo, Y., 2015. Study of hip fracture risk by DXA-based patient-specific finite element model. *Biomed. Mater. Eng.* 25, 213–220.
- Garcia, D., Zysset, P.K., Charlebois, M., Curnier, A., 2009. A three-dimensional elastic plastic damage constitutive law for bone tissue. *Biomech. Model. Mechanobiol.* 8, 149–165.
- Griffith, J.F., Genant, H.K., 2008. Bone mass and architecture determination: state of the art. *Best Pract. Res. Clin. Endocrinol. Metab.* 22, 737–764.
- Hambli, R., Allaoui, S., 2013. A robust 3D finite element simulation of human proximal femur progressive fracture under stance load with experimental validation. *Ann. Biomed. Eng.* 41, 2515–2527.
- Helgason, B., Perilli, E., Schileo, E., Taddei, F., Brynjolfsson, S., Viceconti, M., 2008. Mathematical relationships between bone density and mechanical properties: a literature review. *Clin. Biomech.* 23, 135–146.
- Juszczczyk, M.M., Cristofolini, L., Viceconti, M., 2011. The human proximal femur behaves linearly elastic up to failure under physiological loading conditions. *J. Biomech.* 44, 2259–2266 Epub: June 30 2011.

- Keyak, J.H., Rossi, S.A., Jones, K.A., Skinner, H.B., 1998. Prediction of femoral fracture load using automated finite element modeling. *J. Biomech.* 31, 125–133.
- Keyak, J.H., Sigurdsson, S., Karlsdottir, G.S., Oskarsdottir, D., Sigmarsdottir, A., Kornak, J., Harris, T.B., Sigurdsson, G., Jonsson, B.Y., Siggeirsdottir, K., Eiriksdottir, G., Gudnason, V., Lang, T.F., 2013. Effect of finite element model loading condition on fracture risk assessment in men and women: the AGES-Reykjavik study. *Bone* 57, 18–29.
- Koivumaki, J.E., Thevenot, J., Pulkkinen, P., Kuhn, V., Link, T.M., Eckstein, F., Jamsa, T., 2012. Ct-based finite element models can be used to estimate experimentally measured failure loads in the proximal femur. *Bone* 50, 824–829 Epub: January 28 2012.
- Kopperdahl, D.L., Aspelund, T., Hoffmann, P.F., Sigurdsson, S., Siggeirsdottir, K., Harris, T.B., Gudnason, V., Keaveny, T.M., 2014. Assessment of incident spine and hip fractures in women and men using finite element analysis of CT scans. *J. Bone Miner. Res.* 29, 570–580.
- Larsson, D., Luisier, B., Kersh, M.E., Dall'ara, E., Zysset, P.K., Pandey, M.G., Pahr, D.H., 2014. Assessment of transverse isotropy in clinical-level CT images of trabecular bone using the gradient structure tensor. *Ann. Biomed. Eng.* 42, 950–959.
- Lin, L.I., 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45, 255–268.
- Luisier, B., Dall'ara, E., Pahr, D.H., 2014. Orthotropic HR-pQCT-based FE models improve strength predictions for stance but not for side-way fall loading compared to isotropic QCT-based FE models of human femurs. *J. Mech. Behav. Biomed. Mater.* 32, 287–299.
- Mnif, H., Koubaa, M., Zrig, M., Trabelsi, R., Abid, A., 2009. Elderly patient's mortality and morbidity following trochanteric fracture. A prospective study of 100 cases. *Orthop. Surg. Res.* 95, 505–510 Epub: September 26 2009.
- Morgan, E.F., Bayraktar, H.H., Keaveny, T.M., 2003. Trabecular bone modulus–density relationships depend on anatomic site. *J. Biomech.* 36, 897–904.
- Morgan, E.F., Keaveny, T.M., 2001. Dependence of yield strain of human trabecular bone on anatomic site. *J. Biomech.* 34, 569–577.
- Naylor, K.E., McCloskey, E.V., Eastell, R., Yang, L., 2013. Use of DXA-based finite element analysis of the proximal femur in a longitudinal study of hip fracture. *J. Bone Miner. Res.* 28, 1014–1021.
- Ohman, C., Baleani, M., Pani, C., Taddei, F., Alberghini, M., Viceconti, M., Manfrini, M., 2011. Compressive behaviour of child and adult cortical bone. *Bone* 49, 769–776 Epub: July 6 2011.
- Orwoll, E.S., Marshall, L.M., Nielson, C.M., Cummings, S.R., Lapidus, J., Cauley, J.A., Ensrud, K., Lane, N., Hoffmann, P.R., Kopperdahl, D.L., Keaveny, T.M., 2009. Finite element analysis of the proximal femur and hip fracture risk in older men. *J. Bone Miner. Res.* 24, 475–483.
- Pahr, D.H., Zysset, P.K., 2009. From high-resolution CT data to finite element models: development of an integrated modular framework. *Comput. Methods Biomech. Biomed. Eng.* 12, 45–57.
- Rincon-Kohli, L., Zysset, P.K., 2009. Multi-axial mechanical properties of human trabecular bone. *Biomech. Model. Mechanobiol.* 8, 195–208.
- Sarkalkan, N., Waarsing, J.H., Bos, P.K., Weinans, H., Zadpoor, A.A., 2014. Statistical shape and appearance models for fast and automated estimation of proximal femur fracture load using 2D finite element models. *J. Biomech.* 47, 3107–3114.
- Schileo, E., Balistreri, L., Grassi, L., Cristofolini, L., Taddei, F., 2014. To what extent can linear finite element models of human femora predict failure under stance and fall loading configurations?. *J. Biomech.* 47, 3531–3538.
- Schileo, E., Dall'ara, E., Taddei, F., Malandrino, A., Schotkamp, T., Baleani, M., Viceconti, M., 2008. An accurate estimation of bone density improves the accuracy of subject-specific finite element models. *J. Biomech.* 41, 2483–2491.
- Stevens, J.P., 1984. Outliers and influential data points in resession analysis. *Psychol. Bull.* 95, 334–344.
- Testi, D., Viceconti, M., Baruffaldi, F., Cappello, A., 1999. Risk of fracture in elderly patients: a new predictive index based on bone mineral density and finite element analysis. *Comput. Methods Prog. Biomed.* 60, 23–33.
- Testi, D., Viceconti, M., Cappello, A., Gnudi, S., 2002. Prediction of hip fracture can be significantly improved by a single biomedical indicator. *Ann. Biomed. Eng.* 30, 801–807.
- Treecce, G.M., Gee, A.H., Tonkin, C., Ewing, S.K., Cawthon, P.M., Black, D.M., Poole, K.E., 2015. Predicting hip fracture type with cortical bone mapping (CBM) in the osteoporotic fractures in men (MrOS) study. *J. Bone Miner. Res.*
- Vaananen, S.P., Grassi, L., Flivik, G., Jurvelin, J.S., Isaksson, H., 2015. Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image. *Med. Image Anal.* 24, 125–134.
- Yang, L., Palermo, L., Black, D.M., Eastell, R., 2014. Prediction of incident hip fracture with the estimated femoral strength by finite element analysis of DXA Scans in the study of osteoporotic fractures. *J. Bone Miner. Res.* 29, 2594–2600.
- Zysset, P.K., Dall'ara, E., Varga, P., Pahr, D.H., 2013. Finite element analysis for prediction of bone strength. *BoneKey Rep.* 2, 386.
- Zysset, P., Pahr, D., Engelke, K., Genant, H.K., McClung, M.R., Kendler, D.L., Recknor, C., Kinzl, M., Schwiedrzik, J., Museyko, O., Wang, A., Libanati, C., 2015. Comparison of proximal femur and vertebral body strength improvements in the FREEDOM trial using an alternative finite element methodology. *Bone* 81, 122–130.
- Zysset, P.K., 2003. A review of morphology-elasticity relationships in human trabecular bone: theories and experiments. *J. Biomech.* 36, 1469–1485.