

Published in final edited form as:

Ann Biomed Eng. 2011 February ; 39(2): . doi:10.1007/s10439-010-0196-y.

Robust QCT/FEA Models of Proximal Femur Stiffness and Fracture Load During a Sideways Fall on the Hip

Dan Dragomir-Daescu^{1,2}, Jorn Op Den Buijs⁶, Sean McEligot¹, Yifei Dai¹, Rachel C. Entwistle¹, Christina Salas¹, L. Joseph Melton III^{2,3}, Kevin E. Bennet¹, Sundeep Khosla^{2,4}, and Shreyasee Amin^{2,3,5}

¹Division of Engineering, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA ²College of Medicine, Mayo Clinic, Rochester, MN, USA ³Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA ⁴Divisions of Endocrinology, Diabetes, Metabolism, and Nutrition, Department of Medicine, Mayo Clinic, Rochester, MN, USA ⁵Division of Rheumatology, Department of Medicine, Mayo Clinic, Rochester, MN, USA ⁶Department of Electrical Engineering, Mathematics, and Computer Science, Twente University, Building Carré, Room 4.430, P.O. Box 217, 7500 AE Enschede, The Netherlands

Abstract

Clinical implementation of quantitative computed tomography-based finite element analysis (QCT/FEA) of proximal femur stiffness and strength to assess the likelihood of proximal femur (hip) fractures requires a unified modeling procedure, consistency in predicting bone mechanical properties, and validation with realistic test data that represent typical hip fractures, specifically, a sideways fall on the hip. We, therefore, used two sets ($n = 9$, each) of cadaveric femora with bone densities varying from normal to osteoporotic to build, refine, and validate a new class of QCT/FEA models for hip fracture under loading conditions that simulate a sideways fall on the hip. Convergence requirements of finite element models of the first set of femora led to the creation of a new meshing strategy and a robust process to model proximal femur geometry and material properties from QCT images. We used a second set of femora to cross-validate the model parameters derived from the first set. Refined models were validated experimentally by fracturing femora using specially designed fixtures, load cells, and high speed video capture. CT image reconstructions of fractured femora were created to classify the fractures. The predicted stiffness (cross-validation $R^2 = 0.87$), fracture load (cross-validation $R^2 = 0.85$), and fracture patterns (83% agreement) correlated well with experimental data.

Keywords

Finite element analysis; Bone strength; Hip fracture; Osteoporosis; Quantitative computed tomography; Proximal femur

INTRODUCTION

Despite the significant morbidity and mortality of proximal femur (hip) fractures in aging populations,^{8,32} accurate non-invasive measurements of proximal femur stiffness and strength are not available yet clinically.¹⁰ Dual-energy X-ray absorptiometry (DXA), the current standard for clinical fracture risk assessment, measures areal bone mineral density (aBMD), which is only moderately correlated with bone strength,²³ does not account for variations in bone geometry, and cannot distinguish the cancellous structure from the femoral cortex, though both contribute significantly to bone strength.^{9,27} By contrast, quantitative computed tomography (QCT) can account for the three-dimensional (3D) proximal femur geometry and bone density distribution.¹² Furthermore, the QCT images can provide input to finite element (FE) models (QCT/FEA) to estimate proximal femur strength and predict stresses and strains throughout the bone related to the location and direction of impact forces, which are important determinants of hip fracture.^{4,19,21,28,31} QCT/FEA methods, therefore, have the potential for non-invasive assessment of fracture load, type, and location.¹⁸

Before potential clinical implementation, however, the robustness and convergence of such models must be determined with respect to (1) *meshing* technique and element *size*, (2) bone tissue *segmentation* from the CT scans, and (3) assignment of *material properties*. In addition, the models need to be *validated* by comparing their predictions to *ex vivo* experimental results^{26,28} and clinical fracture data. To date, most previously reported FE model validation studies investigate stance-like loading configurations,^{3,5,19,31} even though a majority of hip fractures are sustained as a result of a fall.⁴ Furthermore, QCT/FEA methods require *bone damage models* to predict fracture loads, but most previously published models have not used a standard failure criterion. For instance, a recent study³ determined the *fracture load* as the load at which at least one solid element had exceeded a minimum principal strain of 1%. Had sensitivity studies been performed, it is likely that such *ad hoc failure criteria* would have shown strong dependencies on mesh density and quality. A more intuitive approach is to determine the value of the predicted force when the force–displacement curve deviates from linearity.¹⁶ Until now, direct comparisons of *predicted* and *measured* fracture loads based on experimental force–displacement curves have not been used to confirm the validity of these damage models.

The objective of our study was to investigate the convergence and robustness of QCT/FEA models that we recently developed for human proximal femur stiffness and ultimate load, and validate the results using experimental test data. High resolution QCT images of femora were used to produce three-dimensional (3D) bone models with improved representation and smoothness of cortical geometry and integration of cancellous volumetric bone mineral density (vBMD). Subsequently, FE models were generated and solved using stepwise linear (but globally non-linear) FE analyses that incorporated a novel bone damage model. Convergence and sensitivity analyses were performed to understand the model robustness to meshing, segmentation, and assignment of material properties. A new damage criterion was also derived, and the resulting equation was then validated on a separate set of femora. To build and validate these models, cadaveric femora were tested to fracture in a sideways fall loading configuration using a testing fixture instrumented with load cells to measure fracture forces. The fracture events were also recorded using a high speed video camera.

MATERIALS AND METHODS

Femur Specimens and Preparation

Two sets of nine fresh frozen, transplant grade cadaveric femora were obtained from 18 individual donors from the Musculoskeletal Transplant Foundation (MTF) (Edison, NJ).

Each set contained three normal (T -score ≥ -1), three osteopenic (T -score, -1 to -2.4), and three osteoporotic (T -score ≤ -2.5) femora, which were classified by femoral neck aBMD while using rice bags to mimic surrounding soft tissues (GE Lunar iDXA, GE Healthcare Inc., Madison, WI).¹⁵ Total proximal femur aBMD was also measured (Table 1). The first set of femora (*training* set) was used to derive model parameters for a new bone *damage* criterion and to study mesh *convergence* and parameter *sensitivity* of QCT/FEA models. The training set contained femora from seven females and two males (mean age at death, 70.4 ± 10.3 years; range: 62–93 years) with a mean femoral neck aBMD of 0.78 ± 0.19 g/cm² (range: 0.47–1.02 g/cm²). The second set (*validation* set) was used to *cross-validate* the model parameters obtained from the training set. The validation set consisted of six female and three male femora (60.0 ± 5.35 years; range: 51–69 years) with a mean femoral neck aBMD of 0.75 ± 0.17 g/cm² (range: 0.48–0.95 g/cm²). Femora were thawed to room temperature for specimen preparation, DXA and QCT scanning, and mechanical testing. Femora were kept moist during all phases of study. We consistently obtained the proximal 250 mm of each femur, removed the soft tissue, and embedded 100 mm of the distal end of the shaft in a block of curing dental cement (Coltène/Whaledent, Cuyahoga Falls, OH), with the femoral neck internally rotated at 15°.

QCT Scanning

Prepared femora were imaged using QCT before and after fracture using a Siemens Somatom Definition Dual Source CT scanner (Siemens Healthcare, Forchheim, Germany). The scanner was operated at 120 kVp, 216 mAs, 1 s rotation time, and pitch = 1, using the scanner's isotropic ultra high resolution mode (zUHR) with collimation 16×0.6 mm. The volume CT dose index (CTDI_{vol}) was 18.8 mGy. Image reconstructions used a sharp (U70) kernel. In-plane pixel size was 0.30–0.45 mm, depending on the transverse dimensions of individual bones. Slice thickness and increment were always 0.4 mm. An in-house designed *scanning fixture* (acrylic) used to hold the femur was aligned with the CT scanner's built-in laser to exactly replicate the femur positioning in the mechanical *testing fixture* (femur shaft at 10° with the horizontal axis, femur neck internally rotated by 15° with respect to the coronal plane). To ensure that each femur had an exact registration of the QCT/FEA models to the testing position, two aluminum rods were embedded in the *scanning fixture*. The rod locations were used to obtain the coordinates for a rotation axis at the distal end of the fixture. A QCT calibration phantom (Mindways Inc., Austin, TX, USA), containing five solid rods of different reference materials, was placed under each femur for individual conversion of Hounsfield units (HU) to equivalent K₂HPO₄ density ($\rho_{K_2HPO_4}$). We obtained the following equation:

$$\rho_{K_2HPO_4} = -9 \cdot 10^{-3} + 7 \cdot 10^{-4} \cdot HU \quad (1)$$

Similar to published work,³⁰ we assumed that $\rho_{K_2HPO_4}$ was identical to bone ash density (ρ_{ash}).

Fracture Testing

The femora were fractured using an instrumented *testing fixture* mounted on a Mini Bionix mechanical testing system (MTS, Minneapolis, MN) (Fig. 1). The cured cement block holding the femur shaft was clamped in the fixture and rigidly attached to a multi-axis load cell (JR3, Woodland, CA) that recorded three force components and three moment components in the femoral shaft during fracture. Right before testing, the greater trochanter was placed in a shallow, cylindrical aluminum cup filled with dental cement. The femoral head was fitted into another aluminum cup connected to linear cross-bearings attached to the MTS actuator (Fig. 1). The linear bearings permitted very low friction displacement in the

horizontal plane. To mimic knee rotation, the femur was allowed to rotate at point O about the x -axis (Fig. 1). Two single-axis load cells (Transducer Techniques, Temecula, CA, USA) recorded femoral head and trochanter vertical reaction forces. Any horizontal forces due to friction were measured using the six-component load cell at the distal end of the fixture. The friction forces at the trochanter site were considered minimal (i.e., one order of magnitude less than the vertical forces) and neglected in this study. A linear displacement sensor (Novotechnik, Southborough, MA, USA) mounted between the crosshead and the actuator ram of the MTS system measured the ram displacement. All signals were sampled at 6250 Hz. High speed video was captured using an APX-RS camera (Photron Inc., San Diego, CA, USA) with a 50 mm f/1.2 lens (Nikon, Tokyo, Japan). An image resolution of 1024×512 pixels and a sampling frequency of 6000 fps were used. All signals and videos were synchronized at the initiation of testing using a trigger pulse. All femora were tested at a vertical displacement rate of 100 mm/s.⁶

QCT/FEA Models

The QCT/FEA process included QCT scanning and segmentation of each proximal femur, rendering of a solid 3D model, generating the finite element mesh from the 3D model, assigning material properties to each finite element, applying boundary conditions, and solving the model (Fig. 2). Detailed modeling steps are presented in the next four sections. Predictive regression models between (a) experimental and QCT/FEA-predicted stiffness, and (b) experimental and QCT/FEA-predicted ultimate load were established from the training set. The relationships established by the training set were applied to the validation set and the cross-validation coefficients of determination (R^2) were calculated and compared. The predictive ability of QCT/FEA-based models for stiffness and ultimate load was compared to similar aBMD-based models using the method developed by Steiger²⁹ ($p < 0.05$). Lastly, we also examined the differences between the slopes and intercepts of the regression lines separately fit to the training and validation sets ($p < 0.05$).

Image-Based Mesh Generation—We generated 3D FE meshes from the reconstructed QCT images using Mimics (Materialise, Ann Arbor, MI). Uniform meshes with maximum element edge lengths of 5.0 mm (*coarse mesh*), 2.5 mm (*medium quality mesh*), and 1.5 mm (*fine mesh*) were created (Figs. 3a–3c). In addition, a *smart mesh* was generated with maximum element edge lengths of 4.0 mm in the distal shaft, 2.5 mm in a transition region in the shaft, and 1.5 mm in the most proximal region (femoral head, neck, greater trochanter, lesser trochanter, and 20 mm of the shaft distal to the lesser trochanter) (Fig. 3d). The smart mesh was generated in ANSYS ICEM (ANSYS, Canonsburg, PA, USA) using an expansion factor of 1.2, such that elements at the cortex were much smaller than elements in the cancellous compartment. Structural stiffness and ultimate load were calculated for each of the femur meshes in the *training set*.

Segmentation Sensitivity Studies—Two different segmentations were used for each bone in the *training set* to study the sensitivity of FEA results to potential errors in generating the 3D models. Segmentation was first obtained using a threshold value of 500 HU. A full layer of pixels was then added on the outside contours of the first segmentation to generate a second segmentation in order to represent a typical error that could be made by an operator. The two segmentations strategies of each femur were meshed using smart meshes, and the FE predictions were compared.

Material Properties Assignment—The robustness of the QCT/FEA models to changes in allocation of discrete material properties was investigated in the *training set*. The mean HU number of each element was averaged from the values of the contained voxels. Elements were then grouped into either 21 or 42 discrete material bins that approximated the

continuous distribution. Subsequently, ash density (ρ_{ash}), isotropic elastic modulus (E), and yield strain (ε_y) were all calculated based on the HU number at the center of each bin. E was calculated from ρ_{ash} based on a previously published relation between apparent density (ρ_{app} , g/cm³) and E (MPa),²⁴ assuming a ratio $\rho_{\text{ash}}/\rho_{\text{app}}$ of 0.6.^{13,17} The following equation was used:

$$E=14664\rho_{\text{ash}}^{1.49} \quad (2)$$

For each material, a Poisson's coefficient of 0.3 was used.

The following power law was assumed between yield strain ε_y and ρ_{ash} (g cm⁻³):

$$\varepsilon_y=0.0081\rho_{\text{ash}}^{-1.42} \quad (3)$$

The constants in Eq. (3) were determined by a trial and error optimization procedure to improve the agreement between predicted and experimental fracture forces in the *training* set.

Finite Element Modeling—The finite element meshes, with their assigned material properties, were imported into ANSYS Mechanical APDL (ANSYS, Canonsburg, PA, USA). To simulate the rotation about O (Fig. 1), two nodes were added to the mesh to define the rotation axis, with coordinates taken from the centers of the two aluminum rods measured from the QCT scans. These nodes were attached with rigid beam elements to the distal end of the femur. A series of static FE analyses were performed using boundary conditions consistent with a sideways fall on the hip (Fig. 2). Displacements of selected nodes at the trochanter surface were set to zero in the z-direction. A simulated force boundary condition was applied in 100 N increments to the femoral head surface. For both the trochanter and the femoral head, the boundary conditions were applied to groups of nodes such that local “crushing” of the bone model was avoided. Finally, the two nodes defining the axis of rotation were assigned boundary conditions allowing rotation about the x-axis only. To simulate damage to the femur, a yield model was incorporated. After each load step, elements with the von Mises strain (ε_{VM}) exceeding the yield strain (ε_y) were “failed” by assigning a very small Young's modulus (0.01 MPa). The model stiffness was then updated, the load was increased, and the model was solved again until the QCT/FEA load–displacement curve reached a plateau. For the convergence studies, a 3.6 mm head displacement was used to define the ultimate load on the trochanter. This value was the *average* displacement in the training set that corresponded to the experimental *ultimate load* (i.e., maximum load). The *average* displacement corresponding to the ultimate load in the validation set was also 3.6 mm. We calculated the FEA *ultimate load* using two methods: (1) at 3.6 mm displacement, and (2) at displacements corresponding to experimental ultimate loads. The *stiffness* of each bone model was calculated from the slope of the initial, linear portion of the force–displacement curve. The fracture locations in the QCT/FEA models were determined from the distributions of the von Mises strains and failed elements. The accuracy of QCT/FEA models was further investigated by comparing the model-predicted fracture patterns to observed patterns in the post-fracture CT reconstructions and the high speed video images.

RESULTS

Convergence Studies for Stiffness and Ultimate Load

The training set QCT/FEA models were used for convergence studies. There were approximately 60,000 elements in the *coarse* meshes, 500,000 in the *medium* meshes, 600,000 in the *smart* meshes, and 2,000,000 in the *fine* meshes. When further increasing the number of elements beyond that in the *fine* mesh, the results did not change (i.e., the model converged). For each femur, we used the fine mesh results as the standard to compare the results generated with all other meshes. For both stiffness and ultimate load, we calculated the percentage differences between the fine mesh and the coarse, medium, and smart meshes, respectively.

The *coarse* mesh strategy produced errors in osteoporotic femur models as high as 35% for stiffness, and as high as 14% for ultimate load. The *medium* quality mesh strategy produced improved but still not convergent results. The errors in osteoporotic femur models were about 17% for stiffness and about 10% for ultimate load. In contrast, *smart* mesh result errors were less than 5% for stiffness, and less than 3% for ultimate load.

The computational times for *fine* meshes were approximately 1 week when using eight cores in a 32-core SUN X4400 server with 256 GB RAM memory and were considered too long for practical purposes. In contrast, the *smart* meshes which matched the accuracy of the fine meshes had a computing time of about 10 h. As such, only smart meshes were used in the subsequent studies.

Segmentation and Material Properties Assignment Sensitivity Studies

For normal bones, the differences in estimated *stiffness* were only about 6% between the two segmentation strategies, while for osteopenic and osteoporotic bones the differences were less than 3%. Very similar differences were observed in the estimation of *ultimate load* between the two strategies. There were negligible differences (~0%) seen in the estimation for stiffness and ultimate load between the use of 21 vs. 42 discrete material bins in models.

QCT/FEA Model Validation with Experimental Data

All experimental fractures shared some common characteristics (Fig. 4). Owing to high compressive forces developed at about 2 mm of machine ram displacement, the base of the femoral neck surrounded by stronger cortical bone began to penetrate the medial aspect of the greater trochanter, which comprises mostly cancellous bone with a very thin cortex (circled area in Fig. 4b, high speed video image). Crushing and sinking occurred at the medial aspect of the greater trochanter in the vast majority of the femora (16/18 bones). The QCT/FEA models reliably captured these damage accumulation events. The von Mises strain distribution showed that the maximum strains similarly occurred in the medial aspect of the greater trochanter and superior aspect of the femoral neck (circled area in Fig. 4b, von Mises strain). The bones continued to yield at flat or slowly varying loads with more bone damage developing during this period. Immediately before fracture, both the recorded deformation and the model-predicted strains showed a growing area of damage on the superior aspect of the femoral neck (Fig. 4c). Finally, a fatal crack initiated at bone-dependent sites under tensile loading conditions. The start of the final crack was always on the medial side of the femur, usually on the neck or slightly distal to the neck. The femur was completely fractured when the measured moments and forces dropped to zero (Fig. 4d). At the time of fracture, the model-predicted strains always showed a zone of intense damage throughout the femur, very often on the neck or the trochanter, indicative of the particular fracture crack (areas enclosed by an ellipse in Fig. 4d).

The QCT/FEA calculated force–displacement curves were compared to experimentally collected curves. Typical results for the vertical (z) components of the reaction forces at the trochanter are shown in Fig. 5 for a normal bone and an osteoporotic bone. In general the model-generated curves were very accurate approximations of the experimental curves. Excellent agreement was observed between *estimated* and *experimental* bone *stiffness* as demonstrated in Fig. 5 by the initial straight portions of the vertical load at the greater trochanter (F_{Tz}) curve. In addition, the yield model implemented by Eq. (3) produced excellent *ultimate load* estimations.

Table 1 presents a summary of test and FEA results for both sets of femora. The experimental *stiffness* values ranged between 894 and 2427 N/mm. The experimental *ultimate load* values ranged from 1409 to 6179 N.

Stiffness and Ultimate Load Estimation Validation

Linear regression models were calculated from the *training* set using QCT/FEA-estimated stiffness and femoral neck aBMD as predictors for experimental *stiffness* (Figs. 6a, 6b). The coefficient of determination using femoral neck aBMD was $R^2 = 0.67$, while the QCT/FEA model was substantially larger ($R^2 = 0.87$). The cross-validation for the *validation* set was $R^2 = 0.44$ for femoral neck aBMD and $R^2 = 0.87$ for QCT/FEA. Similarly, linear regression models were calculated from the *training* set using QCT/FEA-estimated *ultimate load* and femoral neck aBMD as predictors for experimental *ultimate load* (Figs. 6c, 6d). The coefficient of determination using femoral neck aBMD was $R^2 = 0.69$, while the QCT/FEA model was substantially larger ($R^2 = 0.93$). The cross-validation for the *validation* set was $R^2 = 0.78$ for femoral neck aBMD and $R^2 = 0.85$ for QCT/FEA. We performed similar analyses using total hip aBMD. The results were statistically similar, with femoral neck aBMD having a slightly better coefficient of determination for ultimate load (Table 2).

Using the method developed by Steiger,²⁹ we found that QCT/FEA-based stiffness and ultimate load predictions were statistically better in some, but not all cases. In the training set, QCT/FEA was a significantly better predictor of ultimate load than femoral neck aBMD when ultimate load was determined using both experimental ultimate load *and* the load at 3.6 mm displacement with $p = 0.042$ and $p = 0.024$, respectively. Similarly, QCT/FEA was a significantly better predictor of ultimate load than total hip aBMD when ultimate load was determined using a 3.6 mm displacement ($p = 0.046$) and nearly significant when using the experimental ultimate load displacement ($p = 0.080$). In the validation set, QCT/FEA was nearly significantly better at predicting stiffness when compared with femoral neck and total hip aBMD with $p = 0.058$ and $p = 0.063$, respectively.

The QCT/FEA-based regressions were nearly identical when comparing the training and validation sets (Figs. 6a, 6c). The statistical analysis indicated that the slopes and intercepts for the training and validation QCT/FEA-based regressions were not different with $p \geq 0.541$ (Table 2). The aBMD-based regressions looked different upon visual inspection between training and validation sets (Figs. 6b, 6d), but analyses indicated that they were not statistically different at $p = 0.05$ (Table 2). However, in some instances the slopes and intercepts of the training set were nearly significantly different ($0.05 \leq p < 0.10$) from the slopes and intercepts of the validation set for the aBMD-based regressions (Table 2).

Comparison of QCT/FEA Fracture Predictions with Experimental Fractures

The observed experimental fracture patterns for the femora in the *training* and *validation* sets were classified by an orthopedic surgeon as six basicervical, one transcervical, one subcapital, five intertrochanteric, and five pertrochanteric. The QCT/FEA model-estimated distributions of failed elements and ultimate strains always predicted the compressive

damage site observed experimentally at the superior junction of the femoral neck and greater trochanter. Overall, the predicted fracture patterns also matched the observed fracture locations (good agreement in 15 of 18, fair in 3 of 18) (Figs. 7a, 7b).

For the *training* set, eight of nine fracture patterns were correctly predicted. The remaining prediction was categorized as fair; it was complex and included an initial subcapital crack that traveled longitudinally down the femoral neck, ending at the base of the neck as shown in Fig. 7a (femur number seven). The von Mises strains showed a good prediction of both the subcapital and basal damage, while the failed element distribution pointed to more damage at the base of the femoral neck.

For the *validation* set, the predicted and observed fractures patterns were in good agreement in seven of nine cases. For the two remaining osteoporotic femora (numbers seven and eight in Fig. 7b), the predicted cracks from the failed element distributions von Mises strains were slightly different from the experimental fractures. The predicted crack in bone number seven was more proximal than the experimental crack. Bone number eight was a multi-fragmented fracture. The portion of the crack at the base of the femoral neck was correctly predicted while the lesser trochanter fragment was not predicted at all.

DISCUSSION AND CONCLUSIONS

Eighteen cadaveric femora representative of femoral neck aBMD distributions for older ages²⁰ were used to develop and validate QCT/FEA models of cadaveric proximal femur fractures in a simulated loading scenario of a sideways fall on the hip. The experimentally observed stiffness and ultimate load values were consistent with reported data.^{6,7,20} Resulted fracture patterns were classified by an orthopedic surgeon and were consistent with *in vivo* fractures.

A *training* set of nine femora was used to study the robustness of our QCT/FEA models. A novel meshing strategy was developed (i.e., *smart* meshes), which was shown to yield convergent (mesh-independent) results and to be robust to changes in segmentation and material property assignment strategies. Our QCT/FEA-estimated stiffness and ultimate load values were subsequently fit to experimental fracture test results performed at clinically relevant loading rates to develop regression models. The goodness-of-fit of our QCT/FEA-based models was excellent, and the fracture location prediction was very good. These models were then used to validate the QCT/FEA prediction of *stiffness* and *ultimate load* in a separate set of femora (*validation* set).

The QCT/FEA prediction models were capable of accurately reproducing the experimentally determined proximal femur stiffness, ultimate load, and fracture location for a *validation* set of nine femora. The cross-validation R^2 values were high at 0.87 for stiffness and 0.85 for ultimate load. Conversely, we found aBMD of both the femoral neck and total hip to be a weak predictor of proximal femur stiffness. The training and validation regression lines intersect (Figs. 6b, 6d), indicating instability in the predictive power of femoral neck aBMD. When using femoral neck aBMD to predict stiffness, the cross-validation R^2 decreased to 0.44 in the *validation* set from 0.67 in the *training* set. When using femoral neck aBMD to predict *ultimate load*, the cross-validation R^2 increased to 0.78 in the *validation* set from 0.69 in the *training* set. Therefore, when comparing results between the *training* and the *validation* set the opposite directions of change in cross-validation values for stiffness and ultimate load appear to indicate instability in the predictive power of femoral neck aBMD. We performed similar studies using total hip vBMD, obtained from the QCT scans, as a predictor for experimental stiffness and ultimate load. In this study, the DXA-based results

showed consistently better correlations with the experimental data than vBMD. Thus, vBMD data were not reported.

The simultaneous use of von Mises strains and failed element distributions was an important strength of this study. We achieved excellent fracture location predictions for 15 of the 18 femora tested, and the predictions were rather reasonable in the other three cases. In addition, the yield points on the experimental force–displacement curve and the physical site of initial compressive damage were well predicted by QCT/FEA models for all femora. As observed in our high speed videos the vast majority of fractures started with significant accumulation of compressive damage at the junction of the superior aspect of the femoral neck and the greater trochanter. The fatal crack initiated in tension and propagated from the medial to the lateral side of the femur in agreement with previously published observations.¹¹ All observed fracture patterns were similar to reported data and corresponded to fractures seen commonly *in vivo*.^{14,22} Most of these phases were realistically reproduced by our QCT/FEA simulations. In contrast, aBMD cannot be used to predict fracture location. As such, in agreement with other studies,^{5,16} our QCT/FEA models were more efficient and robust predictors of bone fracture.

To simulate bone fracture, a novel yield criterion was introduced. The ρ_{ash} to ε_y relationship was based on the assumption that denser cortical bone fails at lower strains, whereas more ductile cancellous bone can sustain relatively high strains before failure occurs. Although there is evidence that the yield strain of bone *tissue* does not much depend on density,^{1,25} those studies used high resolution μ CT to map the trabecular architecture, making it possible to differentiate bone and soft tissue at the voxel level. In our clinical QCT scans, voxels likely contain a mixture of bone and soft tissue in the cancellous compartment. To reduce computation time, most finite elements in our meshes contained multiple voxels (~10) with their mineral densities averaged, justifying the use of a density-dependent equation for yield strain. Moreover, the yield strains of our proposed power law were fairly close to experimental values within the range of ash densities representing bone tissue.¹ The proposed relation between yield strain and ash density (Eq. 3) resulted in a strong correlation between experimental and predicted ultimate loads. The somewhat reduced accuracy in fracture pattern predictions observed for some osteoporotic bones was likely due to larger zones of reduced cancellous bone density.

One potential limitation of the current failure criterion is that it did not account for strain direction. Previous experimental work indicates that yield strains for bone in compression may be higher than those for bone under tension.² Incorporation of this distinction in the model may further improve fracture load predictions. Also, further refinement of the failure criterion (better power law coefficients or a different yield strain–density function) may improve the QCT/FEA model post-yield behavior leading to better estimates of fracture forces. A published power law was used for the elastic modulus–density relationship in this study (Eq. 2).²⁴ Although this equation originated from experiments conducted on cancellous bone specimens, it extrapolated well to cortical bone densities for use in subject-specific finite element models.²⁸ However, experimental and extrapolation errors to both lower and higher density ranges may have resulted in stiffness estimation inaccuracies.

Although the coefficients of determination were always better for QCT/FEA-based results when compared to femoral neck aBMD, not all observed differences were statistically significant. This was likely due to small sample sizes within the training and validation sets. These differences would have been significant given a sample size of at least $n = 15$ for each set. However, our sets of femora were as balanced as possible with respect to bone density and their characteristics were representative of older adult populations.

Accurate experimental validation of realistic conditions for a fall on the hip is complicated by the speed of the event. In a previous QCT/FEA study,¹⁸ the models were validated with fracture tests performed using only slow loading rates of 0.5 mm/s, although the rate of bone deformation during a fall is likely much higher than the rates used in quasi-static tests. A loading speed of 100 mm/s was used in a previous study⁶ and was considered to be more representative of bone deformation rates during falls from a standing position. In our tests we adopted the same loading rate and made use of the entire force–displacement curve to validate our stiffness, ultimate load, and fracture patterns.

Patient-specific QCT/FEA models will need to be adapted to consider factors such as lower resolution CT scans and radiation exposure. Historically, clinical CT scans have produced results with larger fields of view which encompass the entire pelvic region and consequently reduce the in-plane image resolution. Our field of view was tightly fit to a single cadaveric femur with soft tissue removed thus increasing our in-plane resolution. Although the in-plane resolution is higher in our scans, the radiation exposure is comparable to clinical CT scans. Routine clinical CT scans of the abdomen, pelvis, or combined regions for a standard sized patient use CTDI_{vol} of 16 to 18 mGy at our institution. The radiation exposures used in this study, CTDI_{vol} = 18.8 mGy, are similar to those for clinical CT scans of the same regions of interest. Some high resolution clinical scans of the abdomen-pelvis area use CTDI_{vol} = 24 mGy for a standard patient size. A CTDI_{vol} of 18.8 mGy is the equivalent to approximately 23 mGy maximum skin exposure. This level of skin exposure is about 88 times lower than what may produce skin erythema, a result of high radiation exposure. Our radiation doses compare very favorably and the resolution, although higher than typical proximal femur CT scanning techniques, is achievable using current generation machines. However, our future work is focused on adapting our QCT/FEA models to currently available clinical QCT scans of the proximal femur for *in vivo* validation.

In conclusion, these results indicate the capability of carefully crafted QCT/FEA models to predict proximal femur stiffness, ultimate load, and fracture pattern in a sideways fall, but will require further clinical validation. Our QCT/FEA models of proximal femur strength hold promise as a non-invasive method for an improved quantitative assessment of hip fracture risk.

Acknowledgments

The authors thank Dr. Mark Bolander, Jim Bronk, Vincent Lambert, Dr. Xiaoliang Qin, Alexander Cong, Mike Burke, Brant Newman, Larry Berglund, Dr. Jodie Christner, Dr. Cynthia McCollough, and Elizabeth Atkinson for their valuable contributions to this study. This study was financially supported by the Grainger Foundation: Grainger Innovation Fund and NIH grant AR027065Z-30S1. The authors thank the Musculoskeletal Transplant Foundation for providing the specimens, and the Opus CT Imaging Resource of Mayo Clinic (NIH construction grant RR018898) for CT imaging of the femora.

References

1. Bayraktar HH, Morgan EF, Niebur GL, Morris GE, Wong EK, Keaveny TM. Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. *J Biomech.* 2004; 37:27–35. [PubMed: 14672565]
2. Bessho M, Ohnishi I, Matsumoto T, Ohashi S, Matsuyama J, Tobita K, Kaneko M, Nakamura K. Prediction of proximal femur strength using a CT-based nonlinear finite element method: differences in predicted fracture load and site with changing load and boundary conditions. *Bone.* 2009; 45:226–231. [PubMed: 19398043]
3. 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; 40:1745–1753. [PubMed: 17034798]

4. Bouxsein, ML. Biomechanics of age-related fractures. In: Marcus, R.; Feldman, D.; Nelson, DA.; Rosen, CJ., editors. Osteoporosis. 3. Boston: Elsevier; 2008. p. 601-623.
5. Cody DD, Gross GJ, Hou FJ, Spencer HJ, Goldstein SA, Fyhrie DP. Femoral strength is better predicted by finite element models than QCT and DXA. *J Biomech.* 1999; 32:1013–1020. [PubMed: 10476839]
6. Courtney AC, Wachtel EF, Myers ER, Hayes WC. Effects of loading rate on strength of the proximal femur. *Calcif Tissue Int.* 1994; 55:53–58. [PubMed: 7922790]
7. Courtney AC, Wachtel EF, Myers ER, Hayes WC. Age-related reductions in the strength of the femur tested in a fall-loading configuration. *J Bone Joint Surg Am.* 1995; 77:387–395. [PubMed: 7890787]
8. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. *Lancet.* 2002; 359:1761–1767. [PubMed: 12049882]
9. Dalle Carbonare L, Giannini S. Bone microarchitecture as an important determinant of bone strength. *J Endocrinol Invest.* 2004; 27:99–105. [PubMed: 15053252]
10. Dawson-Hughes, B.; Lindsay, R.; Khosla, S.; Melton, LJ.; Tosteson, ANA.; Favus, M.; Baim, S. Clinician's Guide to Prevention and Treatment of Osteoporosis. Washington, DC: National Osteoporosis Foundation; 2008.
11. de Bakker PM, Manske SL, Ebacher V, Oxland TR, Crompton PA, Guy P. During sideways falls proximal femur fractures initiate in the superolateral cortex: evidence from high-speed video of simulated fractures. *J Biomech.* 2009; 42:1917–1925. [PubMed: 19524929]
12. Genant HK, Gordon C, Jiang Y, Lang TF, Link TM, Majumdar S. Advanced imaging of bone macro and micro structure. *Bone.* 1999; 25:149–152. [PubMed: 10423042]
13. Goulet RW, Goldstein SA, Ciarelli MJ, Kuhn JL, Brown MB, Feldkamp LA. The relationship between the structural and orthogonal compressive properties of trabecular bone. *J Biomech.* 1994; 27:375–389. [PubMed: 8188719]
14. Heini PF, Franz T, Fankhauser C, Gasser B, Ganz R. Femoroplasty augmentation of mechanical properties in the osteoporotic proximal femurs: a biomechanical investigation of PMMA reinforcement in cadaver bones. *Clin Biomech.* 2004; 19:506–512.
15. Kanis JA, McCloskey EV, Johansson H, Oden A, Melton LJ, Khaltav N. A reference standard for the description of osteoporosis. *Bone.* 2008; 42:467–475. [PubMed: 18180210]
16. Keyak JH. Improved prediction of proximal femoral fracture load using nonlinear finite element models. *Med Eng Phys.* 2001; 23:165–173. [PubMed: 11410381]
17. Keyak JH I, Lee Y, Skinner HB. Correlations between orthogonal mechanical properties and density of trabecular bone: use of different densitometric measures. *J Biomed Mater Res.* 1994; 28:1329–1336. [PubMed: 7829563]
18. Keyak JH, Rossi SA, Jones KA, Les CM, Skinner HB. Prediction of fracture location in the proximal femur using finite element models. *Med Eng Phys.* 2001; 23:657–664. [PubMed: 11755810]
19. Keyak JH, Rossi SA, Jones KA, Skinner HB. Prediction of femoral fracture load using automated finite element modeling. *J Biomech.* 1998; 31:125–133. [PubMed: 9593205]
20. Langton CM, Pisharody S, Keyak JH. Comparison of 3D finite element analysis derived stiffness and BMD to determine the failure load of the excised proximal femur. *Med Eng Phys.* 2009; 31:668–672. [PubMed: 19230742]
21. Lotz JC, Cheal EJ, Hayes WC. Fracture prediction for the proximal femur using finite element models. Part I. Linear analysis. *J Biomech Eng.* 1991; 113:353–360. [PubMed: 1762430]
22. Lotz JC, Cheal EJ, Hayes WC. Stress distributions within the proximal femur during gait and falls: implications for osteoporotic fracture. *Osteoporos Int.* 1995; 5:252–261. [PubMed: 7492864]
23. McCreadie BR, Goldstein SA. Biomechanics of fracture: is bone mineral density sufficient to assess risk? *J Bone Miner Res.* 2000; 15:2305–2308. [PubMed: 11127195]
24. Morgan EF, Bayraktar HH, Keaveny TM. Trabecular bone modulus–density relationships depend on anatomic site. *J Biomech.* 2003; 36:897–904. [PubMed: 12757797]
25. Morgan EF, Bayraktar HH, Yeh OC, Majumdar S, Burghardt A, Keaveny TM. Contribution of intersite variations in architecture to trabecular bone apparent yield strains. *J Biomech.* 2004; 37:1413–1420. [PubMed: 15275849]

26. 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; 24:475–483. [PubMed: 19049327]
27. Pors Nielsen S. The fallacy of BMD: a critical review of the diagnostic use of dual X-ray absorptiometry. *Clin Rheumatol.* 2000; 19:174–183. [PubMed: 10870649]
28. Schileo E, Taddei F, Malandrino A, Cristofolini L, Viceconti M. Subject-specific finite element models can accurately predict strain levels in long bones. *J Biomech.* 2007; 40:2982–2989. [PubMed: 17434172]
29. Steiger JH. Tests for comparing elements of a correlation matrix. *Psychol Bull.* 1980; 87:245–251.
30. Suzuki S, Yamamuro T, Okumura H, Yamamoto I. Quantitative computed tomography: comparative study using different scanners with two calibration phantoms. *Br J Radiol.* 1991; 64:1001–1006. [PubMed: 1742578]
31. Taddei F, Cristofolini L, Martelli S, Gill HS, Viceconti M. Subject-specific finite element models of long bones: an in vitro evaluation of the overall accuracy. *J Biomech.* 2006; 39:2457–2467. [PubMed: 16213507]
32. U.S. Department of Health and Human Services, and Office of the Surgeon General. Bone Health and Osteoporosis: A Report of the Surgeon General. Rockville, MD: U.S. Department of Health and Human Services; 2004. p. 436

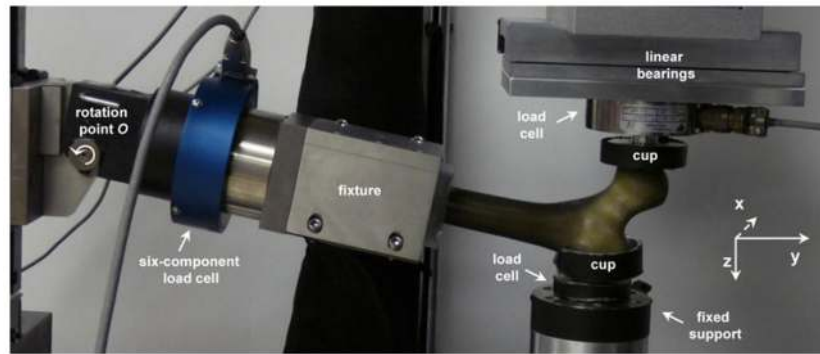


FIGURE 1.

Experimental setup. The distal end of the femur was embedded in a block of dental cement and clamped in a fixture. The fixture was placed at an angle of 10° with the y -axis and could rotate about the x -axis. The fixture contained a six-component load cell. The lateral aspect of the greater trochanter was embedded in a cup with dental cement which was placed on a load cell rigidly attached to the testing machine. The femur head was positioned underneath an aluminum cup. This cup was connected to a load cell that could move in the x and y directions using very low friction linear bearings attached to the machine actuator.

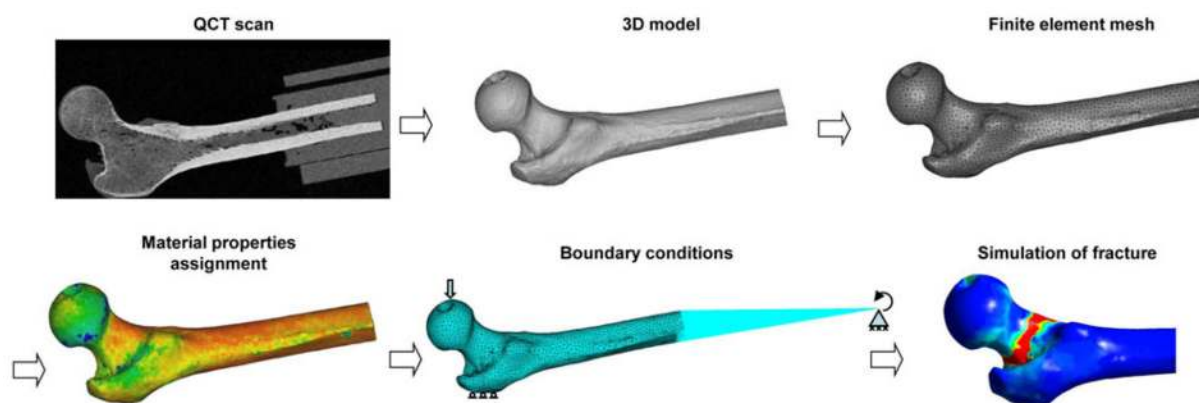


FIGURE 2.

Overview of the QCT/FEA modeling. The QCT scan was segmented to obtain a 3D model. This model was meshed with finite elements. Material properties were assigned to the elements based on the HU values. Boundary conditions were then applied and the model was solved to simulate the fracture.

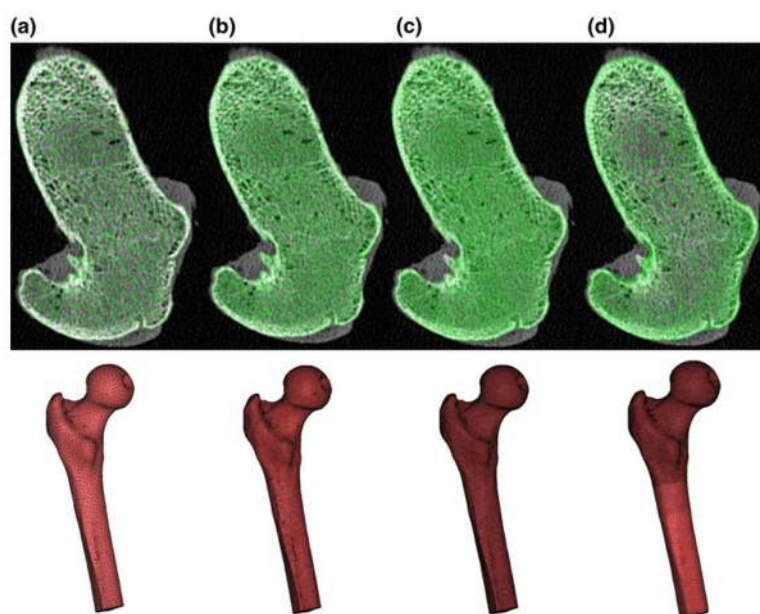


FIGURE 3. Meshes used in the convergence study. Uniform meshes with maximum element edge lengths of (a) 5.0, (b) 2.5, and (c) 1.5 mm were generated. (d) A *smart* mesh with non-uniform element edge lengths was developed to achieve results consistent with the 1.5 mm uniform mesh in a fraction of the computational time.

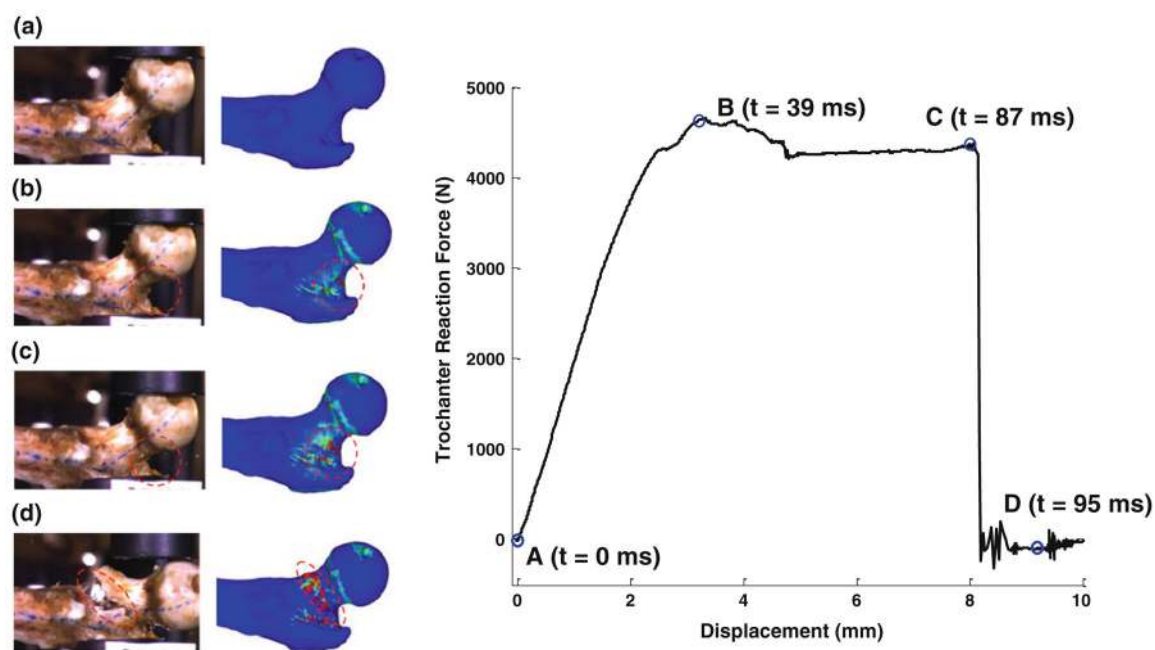


FIGURE 4. Typical load–displacement curve synchronized with fracture events from high speed video. At point a, loading initiates. At point b, ultimate load is reached. Failure occurs at point c. At point d, the bone is completely broken.

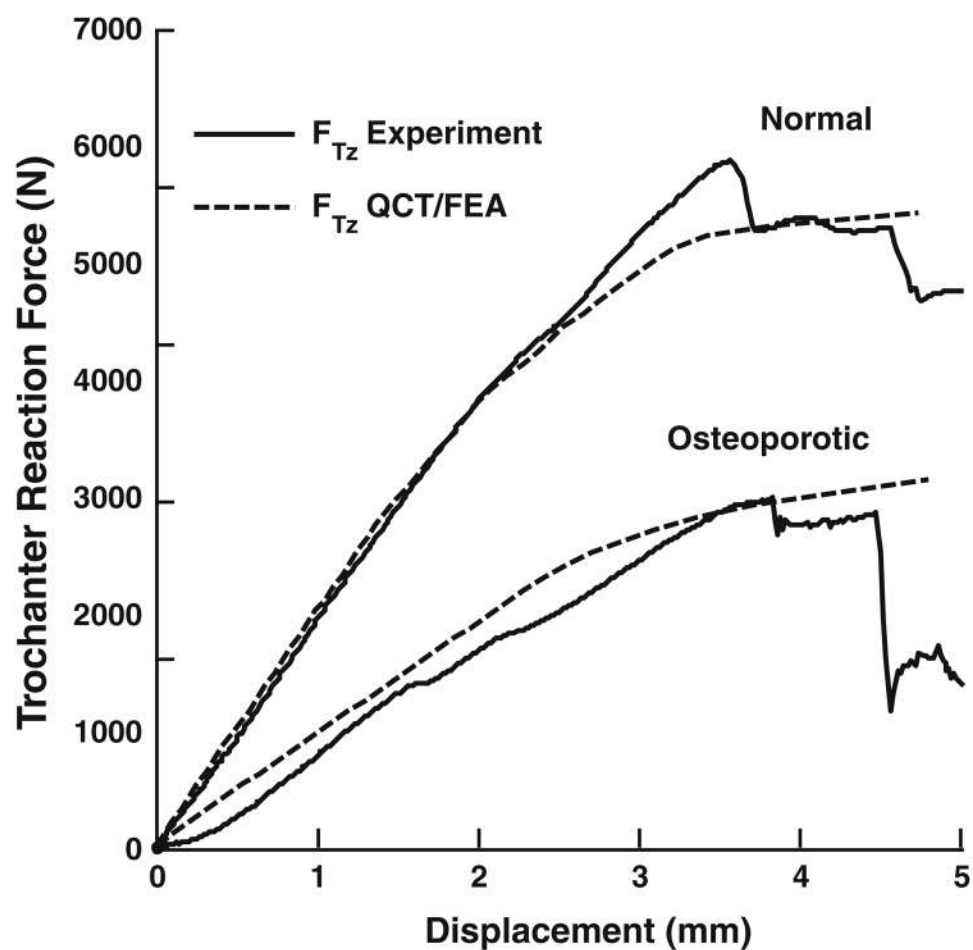


FIGURE 5. Load-displacement curves from experimental data and QCT/FEA models for one normal bone and one osteoporotic bone.

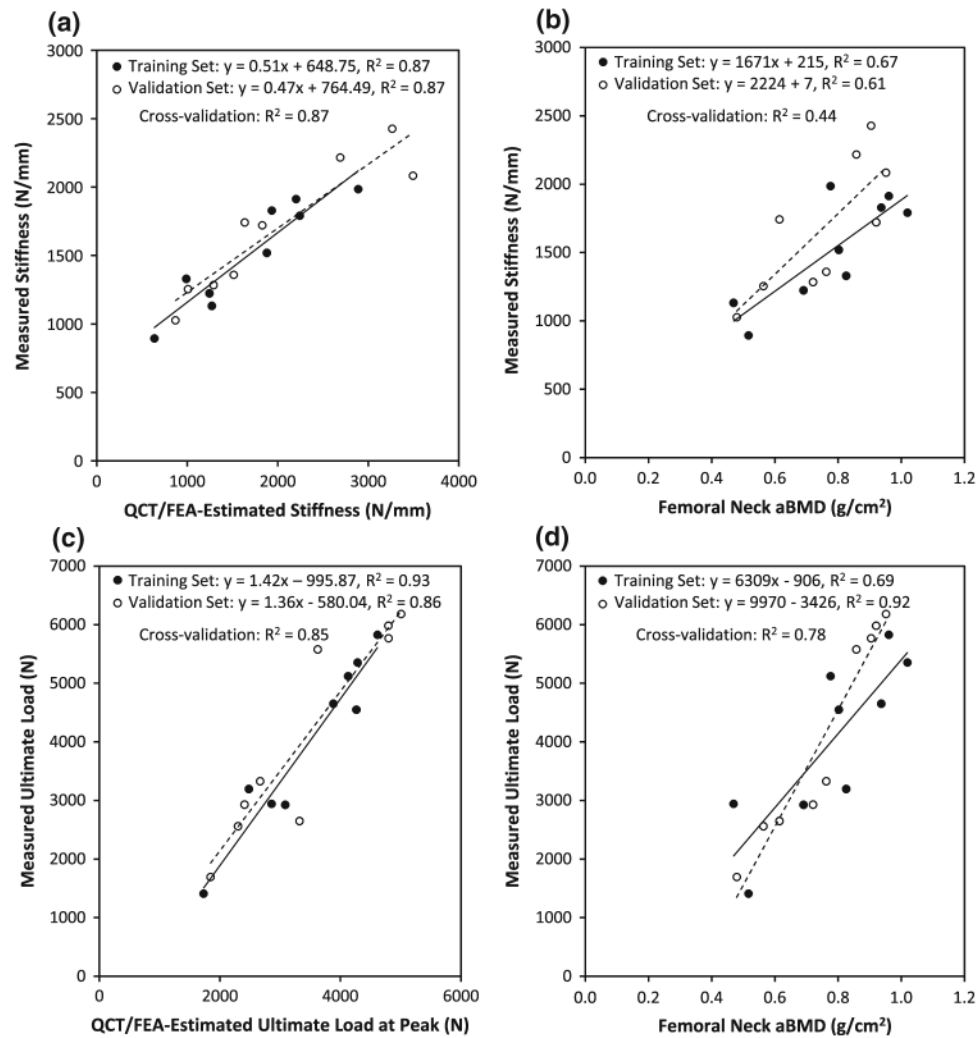

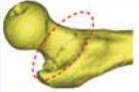
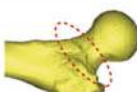

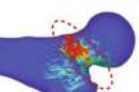




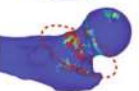

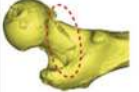


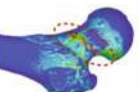

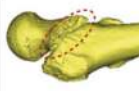


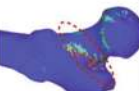




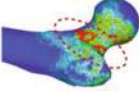




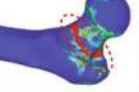

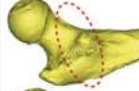
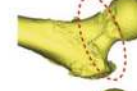

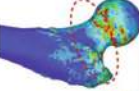




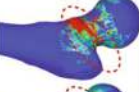




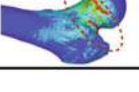


FIGURE 6.

Linear regression models for training and validation sets. Experimental stiffness was predicted by (a) QCT/FEA-estimated stiffness and (b) femoral neck aBMD. Experimental ultimate load was predicted by (c) QCT/FEA-estimated ultimate load and (d) femoral neck aBMD. The models in the training set were cross-validated by the validation set.

(a)

Training Set	High Speed Video	QCT of Fracture View I	QCT of Fracture View II	QCT/FEA Failed Elements	QCT/FEA Von Mises Strain	Degree of Match
Normal						
1						Good
2						Good
3						Good
Osteopenic						
4						Good
5						Good
6						Good
Osteoporotic						
7						Fair
8						Good
9						Good

(b)



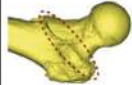

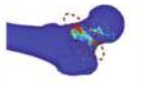




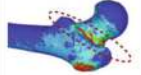




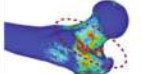

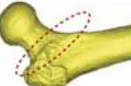


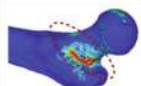




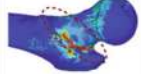




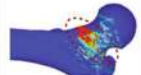

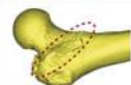


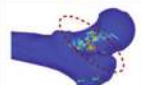




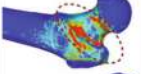





Validation Set	High Speed Video	QCT of Fracture View I	QCT of Fracture View II	QCT/FEA Failed Elements	QCT/FEA Von Mises Strain	Degree of Match
Normal						
1						Good
2						Good
3						Good
Osteopenic						
4						Good
5						Good
6						Good
Osteoporotic						
7						Fair
8						Fair
9						Good

FIGURE 7.
Comparison of fracture pattern predictions to high speed video and CT reconstruction of experimental fractures for (a) the training set and (b) the validation set.

TABLE 1

Details of the 18 investigated proximal femur specimens.

Femur	Age	Gender	Side	Femoral neck aBMD (g/cm ²)	Total hip aBMD (g/cm ²)	Young adult femoral neck T- score	WHO classification	Stiffness (N/mm)		Ultimate load (N)		
								Measured	Predicted	Measured	Predicted at peak	Predicted at 3.6 mm
Training set												
1	63	M	L	0.960	1.082	-0.6	Normal	1913	2202	5825	4615	4638
2	62	F	R	0.936	0.976	-0.7	Normal	1830	1934	4649	3882	3981
3	69	F	L	1.019	1.115	-0.1	Normal	1791	2243	5354	4286	4272
4	62	F	R	0.776	0.995	-1.9	Osteopenic	1986	2889	5120	4129	4052
5	67	F	L	0.825	0.942	-1.5	Osteopenic	1330	989	3194	2481	2515
6	67	F	R	0.802	0.888	-1.7	Osteopenic	1520	1880	4548	4265	3918
7	93	F	L	0.516	0.569	-3.8	Osteoporotic	894	638	1409	1732	1893
8	81	M	L	0.469	0.560	-4.1	Osteoporotic	1132	1272	2941	2859	2939
9	70	F	L	0.690	0.777	-2.5	Osteoporotic	1224	1245	2925	3086	3150
Validation set												
1	56	M	R	0.920	0.951	-0.8	Normal	1721	1830	5980	4798	4677
2	52	F	L	0.951	1.041	-0.6	Normal	2084	3495	6179	5009	4974
3	51	M	R	0.904	0.932	-1.0	Normal	2427	3265	5767	4801	4707
4	65	F	R	0.762	0.813	-2.0	Osteopenic	1359	1514	3327	2667	2673
5	63	F	R	0.720	0.784	-2.3	Osteopenic	1284	1291	2928	2409	2676
6	63	F	L	0.857	0.966	-1.3	Osteopenic	2217	2691	5576	3625	3583
7	67	M	R	0.614	0.676	-3.0	Osteoporotic	1742	1637	2647	3322	3507
8	61	F	R	0.563	0.579	-3.4	Osteoporotic	1255	1010	2559	2302	2515
9	62	F	R	0.479	0.486	-4.0	Osteoporotic	1027	871	1693	1845	2295

Stiffness and ultimate load were measured at the greater trochanter.

TABLE 2

Statistical results for training and validation sets.

Predictor variable	Measured variable	Training set			Validation set			Cross-validation			Slopes		Intercepts	
		Slope	Intercept	R ²	Slope	Intercept	R ²	R ²	RMSE	Different?	p value	Different?	p value	
QCT/FEA stiffness	Experimental stiffness	0.51	648.75	0.87	0.47	764.49	0.87	0.87	198	No	0.701	No	0.599	
Femoral neck aBMD	Experimental stiffness	1671	215	0.67	2224	7	0.61	0.44	413	No	0.495	No	0.126	
Total hip aBMD	Experimental stiffness	1698	22	0.79	2023	55	0.63	0.29	464	No	0.631	Yes	0.024	
QCT/FEA ultimate load at peak	Experimental ultimate load	1.42	-955.87	0.93	1.36	-580.04	0.86	0.85	725	No	0.818	No	0.541	
QCT/FEA ultimate load at 3.6 mm	Experimental ultimate load	1.55	-1401.99	0.94	1.52	-1282.41	0.82	0.82	810	No	0.941	No	0.906	
Femoral neck aBMD	Experimental ultimate load	6309	-906	0.69	9970	-3426	0.92	0.78	892	Almost	0.087	No	0.464	
Total hip aBMD	Experimental ultimate load	6287	-1525	0.78	8775	-2974	0.89	0.71	1022	No	0.172	Almost	0.080	

Results include regression models for each set, cross-validation, and regression model comparisons between sets. RMSE = root mean square error. Significance level of $p = 0.05$.