Title

A new hip fracture risk index derived from FEA-computed proximal femur fracture loads and energies-to-failure.

Permalink

Journal

Authors

Cao, Xuewei
Sigurdsson, Sigurdur
Zhao, Chen
et al.

Publication Date

DOI

A new hip fracture risk index derived from FEA-computed proximal femur fracture loads and energies-to-failure

Xuewei Cao[1], **Joyce H. Keyak**[2], **Sigurdur Sigurdsson**[3], **Chen Zhao**[4], **Weihua Zhou**[4], **Anqi Liu**[5], **Thomas F. Lang**[6], **Hong-Wen Deng**[5], **Vilmundur Gudnason**[3,7], **Qiuying Sha**[1]

[1]Department of Mathematical Sciences, Michigan Technological University, Houghton, MI 49931, USA

[2]Department of Radiological Sciences, Department of Biomedical Engineering, and Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

[3]Icelandic Heart Association Research Institute, Kopavogur, Iceland

[4]Department of Applied Computing, Michigan Technological University, Houghton, MI, USA

[5]Center for Bioinformatics and Genomics, School of Medicine, Tulane University, New Orleans, LA, USA

[6]Department of Radiology and Biomedical Imaging, University of California, San Francisco, CA, USA

[7]University of Iceland, Reykjavik, Iceland

Abstract

Summary—Hip fracture risk assessment is an important but challenging task. Quantitative CT-based patient-specific finite element (FE) analysis (FEA) incorporates bone geometry and bone density in the proximal femur. We developed a global FEA-computed fracture risk index to increase the prediction accuracy of hip fracture incidence.

Purpose—Quantitative CT-based patient-specific finite element (FE) analysis (FEA) incorporates bone geometry and bone density in the proximal femur to compute the force (fracture load) and energy necessary to break the proximal femur in a particular loading condition. The fracture loads and energies-to-failure are individually associated with incident hip fracture, and provide different structural information about the proximal femur.

Methods—We used principal component analysis (PCA) to develop a global FEA-computed fracture risk index that incorporates the FEA-computed yield and ultimate failure loads and energies-to-failure in four loading conditions of 110 hip fracture subjects and 235 age- and sex-matched control subjects from the AGES-Reykjavik study. Using a logistic regression model, we compared the prediction performance for hip fracture based on the stratified resampling.

Conflicts of interest None.

Results—We referred the first principal component (PC1) of the FE parameters as the global FEA-computed fracture risk index, which was the significant predictor of hip fracture (p-value < 0.001). The area under the receiver operating characteristic curve (AUC) using PC1 (0.776) was higher than that using all FE parameters combined (0.737) in the males (p-value < 0.001).

Conclusions—The global FEA-computed fracture risk index increased hip fracture risk prediction accuracy in males.

Keywords

Hip fracture risk; Principal component analysis; Bone strength; Finite element analysis; Osteoporosis

Introduction

Osteoporosis is among the most common and costly metabolic bone diseases [1], causing bones to become weak and brittle and greatly increasing the risk of fracture [2, 3]. Although osteoporosis can affect any bone in the human body, osteoporotic fractures of the proximal femur are the most devastating outcome of the disease, often signaling an end to independent living in the functional elderly. Fracture risk assessment and risk stratification through screening are necessary to reduce the incidence of hip fracture. The areal bone mineral density (aBMD) extracted from the dual-energy X-ray absorptiometry (DXA), a 2D-projection technique, sometimes provides limited information about skeletal factors on fracture risk [4, 5].

Quantitative CT (QCT) imaging is one of the most powerful methods for assessing bone quality in the proximal femur; after three-dimensional (3D) segmentation of the proximal femur, features such as volumetric bone mineral density (BMD) of the cortical and trabecular bone and bone geometry can be computed. Although DXA has undergone many years of development and has benefited from multiple studies of large cohorts and QCT measures have not undergone as much analysis, QCT has the potential to provide improved prediction of hip fracture once additional research and development are performed [6]. QCT is much more sensitive to changes in trabecular BMD and can identify reductions in trabecular BMD well before DXA. DXA aBMD combines cortical and trabecular bone into one measure, thereby obscuring measurements of trabecular BMD and reducing the sensitivity of DXA to changes in trabecular BMD. When bone loss is rapid, such as during use of glucocorticoids or during spaceflight, trabecular bone is lost first, and use of DXA aBMD understates bone loss because cortical bone is superimposed over trabecular bone [7, 8]. Patient-specific finite element analysis (FEA) from QCT images incorporates bone geometry, cortical thickness, and the three-dimensional distribution of bone density in the proximal femur to compute the force (fracture load) necessary to break the proximal femur in a particular loading condition. The fracture load can be defined as the force at the onset of fracture (the yield strength) or the maximum force the proximal femur withstands before complete fracture (the ultimate strength or load capacity), which are calculated using FEA with linear or nonlinear material properties, respectively. FEA-computed fracture loads are the most robust measure of proximal femoral structural integrity and, therefore, hip fracture risk [9, 10]. In particular, the ultimate strength or load capacity of the proximal femur which

is computed with nonlinear FEA is associated with incident hip fracture in men and women, and in men even after controlling for aBMD [11].

Finite element models compute the force necessary to break the proximal femur when forces and other boundary conditions are applied to reflect a specific loading condition, such as single-limb stance during walking or impact onto the greater trochanter from a fall in a specific direction [11]. Although the FEA-computed yield or ultimate strengths for different loading conditions are individually associated with incident hip fracture, and are mutually correlated, each yield and ultimate strength for each loading condition provides different structural information about the proximal femur that contributes to a subject's overall hip fracture risk. The energy transferred to the proximal femur prior to reaching ultimate failure in a particular loading condition may also be associated with fracture risk. Therefore, in this study we used principal component analysis (PCA) to examine a combination of the FEA-computed yield and ultimate strengths and energies-to-failure in various loading conditions to obtain a more robust measure of fracture risk than individual FEA-computed strengths. Principal component analysis is a commonly used method to reduce the dimensionality of data by selecting the most important features that capture the most information [12]. PCA can also speed up the algorithm for prediction by getting rid of correlated variables which do not contribute to decision making [13]. Therefore, we employed PCA to develop a global FEA-computed hip fracture risk index based on the FE model results of 110 hip fracture subjects and 235 age- and sex-matched control subjects in a subset of the AGES-Reykjavik study [14]. PCA is one of the most widely used dimension reduction techniques to transform a large number of variables into a smaller number of variables by identifying the correlations and patterns, while preserving most of the valuable information. The objectives of this study were (1) to construct a global FEA-computed fracture risk index derived from the most important FE parameters of fracture risk based on PCA that is associated with incident hip fracture and (2) to determine if the global FEA-computed fracture risk index, after adjusting for aBMD and covariates, can predict hip fracture better than the FE yield strength, ultimate failure load and energy-to-failure in a single loading condition or their combination in four different loading conditions.

Methods

Subjects

In this study, we used 110 hip fracture subjects and 235 age- and sex-matched control subjects from the Age Gene/ Environment Susceptibility (AGES) Reykjavik study [11, 14]. The AGES-Reykjavik study is an ongoing population-based study which contains the baseline QCT scans of subjects who had no metal implants at the level of the hip [14]. Subjects were followed for 4 to 7 years, through November 15, 2009. Forty-two males and 68 females suffered hip fractures during the follow-up period. Meanwhile, 92 male and 143 female control subjects were selected from a pool of age- and sex-matched subjects. The age of the male fracture group was from 71 to 93 years with a standard deviation (SD) = 5.6 years, and that of the female fracture group was from 67 to 93 years with SD = 5.9 years (Table 1). Accordingly, the age range of the control male group was from 70 to 90 years with SD = 5.2 years, and that of the control female group was from 67 to 92 years with SD =

5.7 years. The average heights were 174.6 cm and 174.9 cm for male fracture and control subjects, respectively. The average heights were 159.6 cm and 159.2 cm for female fracture and control subjects, respectively. In this study, we only included the subjects without any missing data for all covariates and FEA parameters, which is the minor differences in the number of included subjects in the dataset used in Keyak et al. [11].

FEA-computed parameters

The FE models simulated mechanical testing of the femur in which displacement was incrementally applied to the femoral head [11, 15]. The computed reaction force on the femoral head initially increased, reached a peak value (the fracture load), and then decreased. To achieve this mechanical behavior, the FE models employed heterogeneous isotropic elastic moduli, yield strengths, and nonlinear post-yield properties. These properties were computed from the calibrated QCT density (ρ_{CHA}, g/cm^3) of each voxel in an element, which was then used to compute the ash density (ρ_{ash}, g/cm^3) ($\rho_{ash} = 0.0633 + 0.887 \, \rho_{CHA}$), and ρ_{ash} was used to compute mechanical properties [16-18]. Each linear hexahedral finite element measured 3 mm on a side, and the mechanical properties of the element were computed by averaging the values of each property over all voxels in the element, while accounting for the volume fraction of each voxel within the element. Together, these mechanical properties described an idealized density-dependent nonlinear stress–strain curve for each element [11, 15]. Material yield was defined to occur when the von Mises stress exceeded the yield strength of the element. After yield, the plastic flow was modeled assuming a plastic strain-rate vector normal to the von Mises yield surface and isotropic hardening/softening. Displacement was applied incrementally to the femoral head, and the reaction force on the femoral head was computed at each increment as the distal end of the model was fully constrained. For the fall models, the surface of the greater trochanter opposite the loaded surface of the femoral head was constrained in the direction of the displacements while allowing motion transversely.

Based on our FE modeling method from the QCT data [11, 15, 19, 20], twelve FE parameters were evaluated for each subject: yield strength (force at onset of fracture) calculated during single-limb stance (Sy) and impact from a fall onto the posterior (Py), posterolateral (PLy), and lateral (Ly) aspects of the greater trochanter; ultimate strength (failure load capacity) calculated during single-limb stance (Su) and impact from a fall onto the posterior (Pu), posterolateral (PLu), and lateral (Lu) aspects of the greater trochanter; energy-to-failure calculated during single-limb stance (Senergy) and impact from a fall onto the posterior (Penergy), posterolateral (PLenergy), and lateral (Lenergy) aspects of the greater trochanter. The yield strength was defined as the load at which the von Mises stress in 15 contiguous finite elements exceeded the yield strength for the element. Ultimate failure load was defined the maximum FE-computed force on the femoral head. Energy-to-failure was defined as the area under the force versus displacement curve up to the ultimate failure load. The Pearson correlation coefficient was applied to measure the linear correlation among 12 FE parameters. Although fracture parameters were inherently correlated, the fracture load for each loading condition provided different structural information about the proximal femur that contributes to a subject's overall fracture risk. The yield strength was

defined as the load at which the von Mises stress in 15 contiguous finite elements exceeded the yield strength for the element.

Statistical analysis

The data from both male and female subjects were collectively analyzed to examine the overall effects of FE parameters on hip fracture. Additionally, we analyzed male and female subjects separately to assess the gender-specific differences in the effects of hip fracture. In this study, we conduct the following three types of statistical analyses: (i) multiple linear regression analysis was performed to identify which FE parameters were the most important determinants of fracture risk; (ii) principal component analysis was utilized to construct a global FEA-computed risk index based on the FE parameters; (iii) we applied several statistical models to evaluate the hip fracture prediction performance using both the global FEA-computed fracture risk index and individual FE parameters.

(i) Multiple linear regression analysis—To identify which FE parameters were the most important determinants of fracture risk, multiple linear regression analysis was performed with each FE parameter serving as the dependent variable [11]. Based on our previous study [11], fracture status and the demographic parameters, age, sex, height, and weight were considered as candidate independent variables. All FE parameters, height, and weight are standard normalized using the z-score normalization algorithm. To select the most important independent variables in the multiple linear regression, we first performed a simple linear regression model to test the association between each of the candidate independent variables and each of the 12 FE parameters. In the multiple linear regression, we only retained the independent variables with p-value < 0.1. Interactions between fracture status and demographic parameters were also considered as independent variables. Also, if an interaction term was retained, the individual independent variables making up that interaction were retained, regardless of the p-value for the individual independent variable. The multiple linear regression analyses were performed for each of the FE parameters accounting for the retained independent variables. In all of these analyses, FE parameters, age, height, and weight were standardized by subtracting the mean and dividing by the SD of the pooled data.

(ii) Principal component analysis—The FE parameters for each loading condition provided different structural information about the proximal femur that contribute to a subject's overall fracture risk. Therefore, to obtain a more robust measure of fracture risk, we investigated using principal component analysis (PCA) to develop a global FEA-computed risk index based on the FE parameters which were mutually correlated. In addition to analyzing data from males and females combined (the whole sample), we also applied PCA to the male sample and female sample, separately. We used a logistic regression model to test the association between hip fracture status and each of the principal components. Let FX be the fracture status, where $FX = 1$ if the subject suffered a hip fracture and $FX = 0$ otherwise. For the jth principal component, the logistic regression model was expressed by $\text{logit}(Pr(FX)) = \beta_{0j} + \beta_{1j}PC_j$. The top principal components were retained with p-value < 0.05.

(iii) Statistical model for hip fracture prediction—To evaluate the hip fracture prediction performance using the global FEA-computed fracture risk index, we considered six covariates which contain four demographic parameters, age, sex, height, and weight, and two clinical parameters, health status (HEALSTAT; excellent, very good, good, fair, or poor) and bone medication status (BMDMED; yes or no). Meanwhile, we also consider the CT-derived total femur areal bone mineral density (aBMD$_{CT}$) which was calculated from the existing QCT scans of the AGES-Reykjavik cohort. In our previous study, aBMD$_{CT}$ has been reported to be correlated with DXA total femur aBMD with $r = 0.935$ [9]. The area under the receiver operating characteristic (ROC) curve (AUC) indicated the predictive performance for each of the classification models, where the ROC curve showed the relationship between true positive rate and false positive rate. A larger AUC indicated better the performance of the model at distinguishing between the positive and negative classes. We divided the data based on the incidence of hip fracture into a training set (80% of subjects) and test set (20% of subjects) and analyzed male subjects separately from female subjects as well as analyzing the whole sample. To choose the best predictive model based on the training set, we used three linear classification models, namely, logistic regression (logistic), linear discriminant analysis (LDA), and partial least squares analysis (PLS), along with six non-linear classification models, random forest (RF), quadratic discriminant analysis (QDA), mixture discriminant analysis (MDA), neural networks (NNET), multivariate adaptive regression splines (MARS), and K-nearest neighbors (KNN) [21]. In the training set, we used stratified leave-one-group-out cross-validations (LGOCVs), repeating this procedure 25 times. For each LGOCV, we used 75% of the data to build the classification models and 25% of the data to predict and calculate the AUC. After choosing the best models, we built those models using the entire training set and predicted the fracture status based on the test set. To compare the performance for predicting hip fracture using (a) the global FEA-computed fracture risk index, aBMD$_{CT}$, and covariates, (b) the FE parameters, aBMD$_{CT}$, and covariates, and (c) aBMD$_{CT}$ and covariates, we performed stratified resampling 1000 times. Then, we applied a one-sided Student's t-test to compare the resampled AUCs since we were interested in whether the predicative performance of one model was significantly better than the other model. To assess the disparity in predictive performance for hip fractures using the above three resampled AUC values, we also utilized the one-way analysis of variance (ANOVA).

Results

Within the male sample, the fracture and control groups were not significantly different with respect to age, height, weight, and the three energies-to-failure in the fall loading conditions (PLenergy, Penergy, Lenergy) at the time of the CT scan (p-value > 0.120; Table 1). However, the remaining nine FE parameters were significantly lower in each fracture group than in the respective control group (p-value < 0.001). In contrast, within the female sample, the fracture and control groups were significantly different with respect to weight (p-value = 0.036). As for the male group, age, height, and the three energies-to-failure in the fall loading conditions in the female group (PLenergy, Penergy, Lenergy) were not significantly different between fracture and control groups (p-value > 0.513). Similarly to the male group, the remaining nine FE parameters were also significantly lower in each

fracture group than in the respective control group (Lu, p-value = 0.007; all others, p-value < 0.001; Table 1).

For the multiple linear regression analyses (Table 2), after controlling for four demographic parameters (age, sex, height, and weight) and interactions, the FE parameters except for Pu, and energies-to-failure in the fall loading conditions were associated with hip fracture (p-value < 0.1). However, Pu was significantly lower in each fracture group than in the respective control group when not controlling for demographic parameters and interactions (p-value < 0.001) (Table 1) and had the higher R^2 = 0.5217 (Table 2). Note that the interaction of fracture and weight was not retained in our analysis since all p-values for this interaction were greater than 0.1. Therefore, we considered nine hip fracture-related FE parameters (Sy, Su, Senergy, Py, Pu, PLy, PLu, Ly, and Lu) in the following analysis. PCA was applied to these nine FE parameters, which were inherently highly correlated. The proportions of variance explained by the first principal component (PC1) were 83.31%, 78.23%, and 80.65% for the whole sample, male sample, and female sample, respectively. We found that PC1 of the FE parameters was the only significant predictor for hip fracture (p-value < 0.001 in the whole sample, male sample, and female sample). Therefore, we referred to PC1 as the global FEA computed fracture risk index. Using the LGOCV, we found that the performance of using PC1 along with aBMD$_{CT}$ and covariates, or the nine FE parameters combined along with aBMD$_{CT}$ and covariates, were better than that of only using aBMD$_{CT}$ and covariates to predict hip fracture (Table 3). In particular, we observed the superior predictive performance within the whole sample and male sample of PC1 compared with FE parameters combined for all nine classification models; however, within the female sample, FE parameters combined had greater AUC than PC1. Logistic regression and PLS had the greatest AUCs among nine models (Table 3), and were chosen as the two best models. For PLS and logistic based on stratified resampling (Figs. 1 and 2), in the whole sample, logistic performed better than the PLS by using PC1 and FE parameters combined along with aBMD$_{CT}$ and covariates, which were also better than Logistic using aBMD$_{CT}$ and covariates (p-value < 0.001). In contrast, within the male sample, PLS using PC1 and FE parameters combined with aBMD$_{CT}$ and covariates was better than PLS using aBMD$_{CT}$ and covariates (p-value < 0.001).

We performed ANOVA for AUC based on logistic. We observed that the resampled AUCs with aBMD$_{CT}$ and covariates were different with the resampled AUCs with PC1 or FE parameters combined in the whole sample (p-value = 0.05) and in the male sample (p-value = 0.01), but no such difference was found in the female sample (p-value = 0.68). According to ANOVA for AUC based on PLS, the difference was still observed in the male sample with p-value < 0.01. Remarkably, within the female sample, a distinct pattern emerged from PLS, which is consistent with the Student's t-test results shown in Fig. 1. There were different resampled AUCs in the female sample with p-value = 0.05. In contrast, in the whole sample, there was no difference observed (p-value = 0.97).

In particular, we observed the superior predictive performance in male sample of PC1 compared with the predictive performance of each FE parameter based on PLS (p-value = 0.02 for Ly; all others, p-value < 0.01) (Fig. 2). For the female sample, all of the AUCs of PLS using one of the FE parameters were greater than the corresponding AUC using PC1

with p-value > 0.1, except for Ly (p-value = 0.0145). We also observed similar predictive performance based on logistic. In the male sample, the predictive performance using PC1 was better than the predictive performance of each FE parameter (p-value = 0.02 for Su; p-value = 0.04 for Ly; all others, p-value < 0.01). In the female sample, all of the AUCs of logistic using one of the FE parameters were similar to the corresponding AUC using PC1 with p-value > 0.1, except for Ly (p-value < 0.01) and Senergy (p-value = 0.04).

Discussion

This is the first study to construct a global FEA-computed risk index by principal component analysis based on multiple fracture-related FEA-computed fracture loads and energies under different loading conditions. The global FEA-computed fracture risk index, after adjusting for aBMD$_{CT}$ and covariates, predicted hip fracture better than each individual FE parameter (yield strength, ultimate failure load, and energy-to-failure) and also better than the FE parameters combined in the whole sample and the male sample but not the female sample (Figs. 1 and 2; Table 3). Meanwhile, predicting fracture in the female sample was inherently more difficult than predicting fracture in the male sample because the difference between female fracture and control subjects is much smaller than that for males. In our previous work [11], the age-matched design of this study enhanced our ability to explore gender differences in proximal femoral strength and incident hip fracture as a function of age. In particular, our cross-sectional analysis of age-related FE parameter (Su, PLu, Pu, and Lu) loss by gender and fracture status may explain why proximal femoral strength was strongly associated with incident hip fracture in men but much less so in women (Table 2).

Although the FE parameters are individually associated with incident hip fracture, and are mutually highly correlated, they each provide different structural information about the proximal femur that can influence a subject's overall fracture risk. The superior performance, in both men and women, of the assessment of hip fracture risk by using the global FEA-computed risk index and FE parameters combined along with aBMD$_{CT}$ and covariates compared with aBMD$_{CT}$ and covariates is not surprising. The CT-derived total femur aBMD was considered, which has a strong correlation with aBMD from DXA (r = 0.935) [9]. The predictive performance by incorporating information from FE parameters was better than only using aBMD$_{CT}$ and covariates in the male sample and female sample (Fig. 1; Table 3), implying that the FE parameters can provide additional information in the assessment of hip fracture risk by incorporating bone geometry, cortical thickness, and the three-dimensional distribution of bone density in the proximal femur. Meanwhile, the global FEA-computed risk index plus the aBMD from DXA was better than the aBMD only (Fig. 1; Table 3). Although aBMD from DXA correlated with bone weakness and fragility fracture [22], DXA is a 2D-projection technique that poorly accounts for bone geometry and size and is a poor predictor of hip fracture in subjects with osteopenia (T-scores between −1 and −2.5). Thus, aBMD provides limited information about skeletal factors on fracture risk. However, several studies have been conducted using 2D-derived indices other than aBMD which showed better and independent predictive capacity, such as the hip structure analysis techniques [23], the 2D FEA [24], and the hip axis length [25]. With the emergence of QCT scan-based FE modeling, better estimates of proximal femoral strength have become possible [26]. Our study employed three-dimensional (3D)

FE models from QCT scans, which explicitly represent the 3D geometry and distribution of material properties that make each femur structurally and mechanically unique and is therefore more robust than two-dimensional (2D) models from DXA. Principles of physics dictate that hip fracture occurs when an excessive force is applied to the proximal femur, i.e., when the applied force exceeds the force that the proximal femur can support. This force, which varies with the type of loading and force direction, is known as the proximal femoral strength, fracture load [11, 20], or load capacity [15] computed using patient-specific FEA. Patient-specific FEA-computed fracture loads and energies are the most robust measures of proximal femoral structural integrity and, therefore, benefitted for evaluating the hip fracture risk. The comparison of QCT FEA-computed fracture loads and energies and other measurements derived from 2D images deserves further study.

Conclusions

In summary, FEA-computed fracture loads and energies were associated with incident hip fracture in most of loading conditions that were examined, and the global FEA-computed fracture risk index that was investigated by principal component analysis increased hip fracture risk prediction accuracy in the male sample more than that in the female sample. The global FEA-computed fracture risk index was most strongly associated with incident hip fracture in men after accounting for aBMD from DXA and other clinical and demographic parameters. Thus, FE parameters from 3D FE models includes information about hip fracture beyond that of aBMD from 2D models, especially in the male sample. The significance and complexity of these findings, particularly with respect to sex and age effects, indicate that additional studies of FE modeling for hip fracture risk assessment are likely to enhance our understanding of this significant public health problem.

Acknowledgements

Data Availability

The datasets analyzed during the current study are not publicly available. Requests for access to these datasets should be directed to the corresponding authors.

References

1. Richards JB, Zheng H-F, Spector TD (2012) Genetics of osteoporosis from genome-wide association studies: advances and challenges. Nat Rev Genet 13:576–588 [PubMed: 22805710]

2. Recker R, Kimmel D (1991) Changes in trabecular microstructure in osteoporosis occur with normal bone remodeling dynamics. J Bone Miner Res 6:S225

3. Yang T-L, Chen X-D, Guo Y, Lei S-F, Wang J-T, Zhou Q, Pan F, Chen Y, Zhang Z-X, Dong S-S (2008) Genome-wide copy-number-variation study identified a susceptibility gene, UGT2B17, for osteoporosis. The American Journal of Human Genetics 83:663–674 [PubMed: 18992858]

4. Genant H, Engelke K, Prevrhal S (2008) Advanced CT bone imaging in osteoporosis. Rheumatology 47:iv9–iv16 [PubMed: 18556648]

5. Chang G, Honig S, Brown R, Deniz CM, Egol KA, Babb JS, Regatte RR, Rajapakse CS (2014) Finite element analysis applied to 3-T MR imaging of proximal femur microarchitecture: lower bone strength in patients with fragility fractures compared with control subjects. Radiology 272:464 [PubMed: 24689884]

6. Carballido-Gamio J, Bonaretti S, Saeed I, Harnish R, Recker R, Burghardt AJ, Keyak JH, Harris T, Khosla S, Lang TF (2015) Automatic multi-parametric quantification of the proximal femur with quantitative computed tomography. Quant Imaging Med Surg 5:552 [PubMed: 26435919]

7. Keyak J, Koyama A, LeBlanc A, Lu Y, Lang T (2009) Reduction in proximal femoral strength due to long-duration spaceflight. Bone 44:449–453 [PubMed: 19100348]

8. Sibonga J, Spector E, Keyak J, Zwart S, Smith S, Lang T (2020) Use of quantitative computed tomography to assess for clinically-relevant skeletal effects of prolonged spaceflight on astronaut hips. J Clin Densitom 23(2):155–164 [PubMed: 31558405]

9. Keyak J, Sigurdsson S, Karlsdottir G, Oskarsdottir D, Sigmarsdottir A, Zhao S, Kornak J, Harris T, Sigurdsson G, Jonsson B (2011) Male–female differences in the association between incident hip fracture and proximal femoral strength: a finite element analysis study. Bone 48:1239–1245 [PubMed: 21419886]

10. Orwoll ES, Marshall LM, Nielson CM, Cummings SR, Lapidus J, Cauley JA, Ensrud K, Lane N, Hoffmann PR, Kopperdahl DL (2009) Finite element analysis of the proximal femur and hip fracture risk in older men. J Bone Miner Res 24:475–483 [PubMed: 19049327]

11. Keyak J, Sigurdsson S, Karlsdottir G, Oskarsdottir D, Sigmarsdottir A, Kornak J, Harris T, Sigurdsson G, Jonsson B, Siggeirsdottir K (2013) Effect of finite element model loading condition on fracture risk assessment in men and women: the AGES-Reykjavik study. Bone 57:18–29 [PubMed: 23907032]

12. Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. Phil Trans R Soc A 374:20150202 [PubMed: 26953178]

13. Sarker IH (2021) Machine learning: algorithms, real-world applications and research directions. SN Computer Science 2:1–21

14. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF (2007) Age, gene/environment susceptibility–Reykjavik study: multidisciplinary applied phenomics. Am J Epidemiol 165:1076–1087 [PubMed: 17351290]

15. Keyak J, Kaneko T, Khosla S, Amin S, Atkinson E, Lang T, Sibonga J (2020) Hip load capacity and yield load in men and women of all ages. Bone 137:115321 [PubMed: 32184195]

16. Keyak JH, Kaneko TS, Tehranzadeh J, Skinner HB (2005) Predicting proximal femoral strength using structural engineering models. Clin Orthop Relat Res 1976–2007(437):219–228

17. Kaneko T, Bell J, Pejcic M, Tehranzadeh J, Keyak J (2004) Mechanical properties, density and quantitative CT scan data of trabecular bone with and without metastases. J Biomech 37(4):523–530 [PubMed: 14996564]

18. Kaneko T, Pejcic M, Tehranzadeh J, Keyak J (2003) Relationships between material properties and CT scan data of cortical bone with and without metastatic lesions. Med Eng Phys 25(6):445–454 [PubMed: 12787982]

19. Keyak JH, Rossi SA, Jones KA, Skinner HB (1997) Prediction of femoral fracture load using automated finite element modeling. J Biomech 31:125–133

20. Keyak JH (2001) Improved prediction of proximal femoral fracture load using nonlinear finite element models. Med Eng Phys 23:165–173 [PubMed: 11410381]

21. Kuhn M, Johnson K (2013) Applied predictive modeling. Springer

22. Marshall D, Johnell O, Wedel H (1996) Meta-analysis of how well measures of bone mineral density predict occurrence of osteoporotic fractures. BMJ 312:1254–1259 [PubMed: 8634613]

23. LaCroix A, Beck T, Cauley J, Lewis C, Bassford T, Jackson R, Wu G, Chen Z (2010) Hip structural geometry and incidence of hip fracture in postmenopausal women: what does it add to conventional bone mineral density? Osteoporos Int 21:919–929 [PubMed: 19756830]

24. Leslie W, Luo Y, Yang S, Goertzen A, Ahmed S, Delubac I, Lix L (2019) Fracture risk indices from DXA-based finite element analysis predict incident fractures independently from FRAX: the Manitoba BMD Registry. J Clin Densitom 22(3):338–345 [PubMed: 30852033]

25. Leslie W, Lix LM, Morin S, Johansson H, Odén A, McCloskey E, Kanis J (2015) Hip axis length is a FRAX- and bone density-independent risk factor for hip fracture in women. J Clin Endocrinol Metab 100(5):2063–2070 [PubMed: 25751108]

26. Cody DD, Gross GJ, Hou FJ, Spencer HJ, Goldstein SA, Fyhrie DP (1999) Femoral strength is better predicted by finite element models than QCT and DXA. J Biomech 32:1013–1020 [PubMed: 10476839]
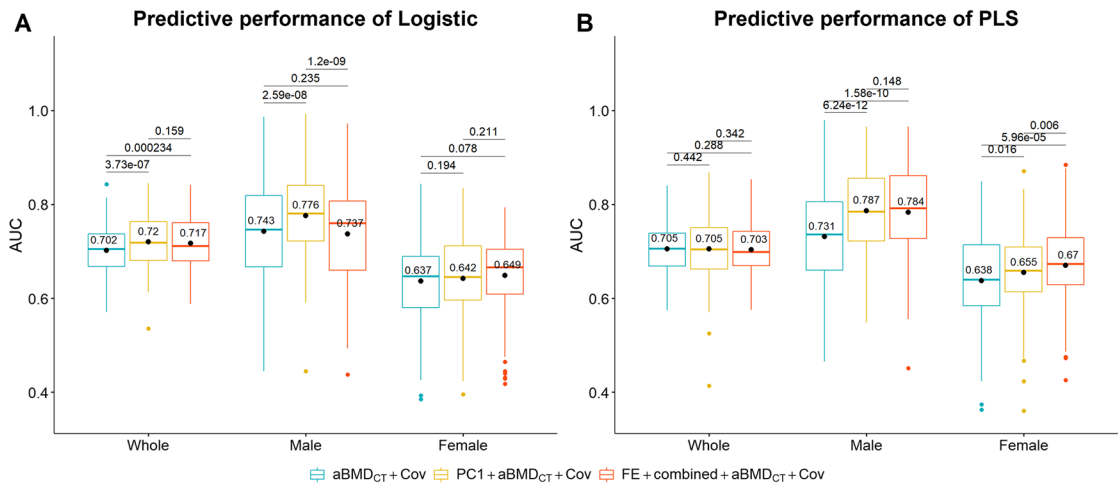
Fig. 1.

The predictive performance of logistic and PLS based on stratified resampling. The values over the lines indicated the *p*-values obtained from the one-sided Student's t-test. The values on top of the boxes indicated the average AUCs by using PC1 along with $aBMD_{CT}$ and covariates ($PC1 + aBMD_{CT} + Cov$) versus using $aBMD_{CT}$ and covariates ($aBMD_{CT} + Cov$) or FE parameters combined with $aBMD_{CT}$ and covariates ($FE combined + aBMD_{CT} + Cov$). Notes: The abbreviations of FE parameters are yield strength (force at onset of fracture) calculated during single-limb stance (Sy) and impact from a fall onto the posterior (Py), posterolateral (PLy), and lateral (Ly) aspects of the greater trochanter; ultimate strength (failure load capacity) calculated during single-limb stance (Su) and impact from a fall onto the posterior (Pu), posterolateral (PLu), and lateral (Lu) aspects of the greater trochanter; energy-to-failure calculated during single-limb stance (Senergy) and impact from a fall onto the posterior (Penergy), posterolateral (PLenergy), and lateral (Lenergy) aspects of the greater trochanter
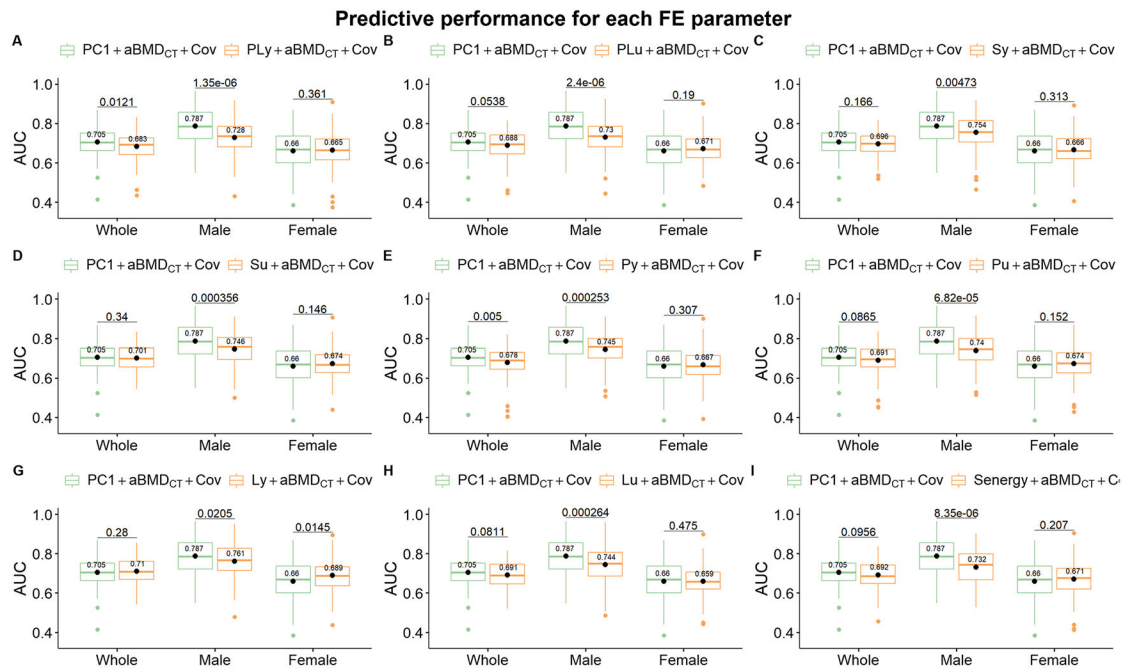
Fig. 2.

The predictive performance of PLS using each FE parameters compared with that of using PC1. The values over the lines indicated the *p*-values were obtained from the one-sided Student's t-test. The values on top of the boxes indicated the average AUCs by using PC1 along with aBMD$_{CT}$ and covariates (PC1 + aBMD$_{CT}$ + Cov) versus using each of FE parameters along with aBMD$_{CT}$ and covariates (* + aBMD$_{CT}$ + Cov)

Table 1

Descriptive statistics for subjects with incident hip fracture and the age- and sex-matched controls. Statistical significance of differences between cases and controls was established using Student's t-test

| Sex	Measure	Controls						Cases						p-value
		N	Min	Max	Median	Average	SD	N	Min	Max	Median	Average	SD	
Male	Age	92	70	90	80	79.7	5.2	42	71	93	81	80.5	5.8	0.432
	Height	92	162.4	190.5	174.2	175	6.6	42	159.2	187.1	174	174.9	5.4	0.941
	Weight	92	52.4	135	82	82.8	14.9	42	50.3	111.3	77.5	78.7	13.5	0.120
	PLy	92	638	4257	1626	1704	673.9	42	276	2254	1202.5	1202.4	465.1	<0.001
	PLu	92	2408	5499	3733.5	3838.6	570.2	42	1723	4269	3253.5	3311.1	529.5	<0.001
	Sy	92	2043	13,990	4697.5	5010.5	1912.5	42	1261	6824	3380.5	3524.4	1219.2	<0.001
	Su	92	4772	21,784	9992.5	10,628.1	2967.8	42	3123	12,898	7980.5	7980.2	2017.2	<0.001
	Py	92	631	4123	1699	1821.3	655.7	42	404	2014	1295	1294.6	415.9	<0.001
	Pu	92	2149	5078	3322	3322.9	456.1	42	1766	3964	3013	2954.9	426.3	<0.001
	Ly	92	1028	5393	2063	2141.5	745.3	42	441	2957	1427	1476.1	507.8	<0.001
	Lu	92	2344	6619	4299.5	4315.2	743.9	42	1913	4794	3776	3669.1	598.1	<0.001
	Senergy	92	287.9	2591	902.9	963.3	425.5	42	189.2	1525.2	641.5	673	282.2	<0.001
	PLenergy	92	233.8	991.9	471.4	496	139.8	42	172.5	1275.1	449.3	483.5	176.4	0.686
	Penergy	92	189.9	788.5	406.4	423.1	137.3	42	218.3	609.7	408.3	420.3	97.7	0.890
	Lenergy	92	89.6	944.2	477.2	475.1	161.8	42	123.8	899.6	440.2	452.2	174.1	0.474
Female	Age	143	67	92	79	79.2	5.7	68	67	93	79.5	79.7	6.1	0.544
	Height	143	139.2	172.9	159.5	159.3	5.6	68	145.1	173.4	159.5	159.3	6.1	0.995
	Weight	143	37.2	112	66.4	68.7	13.7	68	39.2	111.2	63.4	64.2	15	0.036
	PLy	143	314	3071	987	1093.3	476.5	68	353	1945	780	850	336	<0.001
	PLu	143	1816	4535	2837	2963.9	551.2	68	1700	4419	2632	2693.3	440	<0.001
	Sy	143	1272	9446	2924	3362.6	1543.1	68	1300	5913	2528	2680.2	891.6	<0.001
	Su	143	3611	16,330	6775	7256.1	2335.8	68	3085	10,675	5831.5	6045.3	1609.2	<0.001
	Py	143	525	2733	1154	1247.8	468.9	68	490	2049	928	1006.9	341.6	<0.001
	Pu	143	1933	3893	2539	2641.6	399.4	68	1610	3150	2373	2425.3	349	<0.001
	Ly	143	469	3736	1319	1432.2	585.8	68	570	2532	1004	1074.5	382	<0.001
	Lu	143	1941	5386	3203	3256.7	650.6	68	1945	4896	2989.5	3019.9	551.4	0.007

Sex	Measure	Controls							Cases							p-value
		N	Min	Max	Median	Average	SD		N	Min	Max	Median	Average	SD	
	Senergy	143	195.2	1770.3	492.6	558.4	275.8		68	196.5	1115.5	396.7	423.3	155.1	< 0.001
	PLenergy	143	56.1	800.1	403.6	414.6	125.3		68	118.6	912.6	393.2	412.1	139	0.899
	Penergy	143	84.9	787.1	362.9	377	118.9		68	158.1	606.2	360.8	367	96.5	0.513
	Lenergy	143	86.1	943.7	385.8	390.9	154.3		68	119.2	775.4	384.5	382.8	147	0.714

The abbreviations of FE parameters are yield strength (force at onset of fracture) calculated during single-limb stance (Sy) and impact from a fall onto the posterior (Py), posterolateral (PLy), and lateral (Ly) aspects of the greater trochanter; ultimate strength (failure load capacity) calculated during single-limb stance (Su) and impact from a fall onto the posterior (Pu), posterolateral (PLu), and lateral (Lu) aspects of the greater trochanter; energy-to-failure calculated during single-limb stance (Senergy) and impact from a fall onto the posterior (Penergy), posterolateral (PLenergy), and lateral (Lenergy) aspects of the greater trochanter

Table 2

Multiple linear regression results for each of the 12 FE parameters for all subjects, including coefficients of standardized variables, standard errors, and p-values. Fx: fracture status; Fx:Age: the interaction of fracture with age; Fx:Sex: the interaction of fracture with sex; Fx:Height: the interaction of fracture with height; all abbreviations of FE parameters are detailed in the Methods section

Dependent variable	Regression coefficients for standardized variables (Standard error) p-value								R^2
	FX	**Sex**	**Age**	**Weight**	**Height**	**Fx:Age**	**Fx:Sex**	**Fx:Height**	
PLy	**−0.301** (0.115) **0.009**	**0.706** (0.158) **< 0.001**	**−0.147** (0.055) **0.007**	**0.298** (0.056) **< 0.001**	0.033 (0.107) 0.759	-	**−0.420** (0.183) **0.022**	-	0.3773
PLu	**−0.299** (0.101) **0.003**	**0.833** (0.138) **< 0.001**	**−0.151** (0.040) **< 0.001**	**0.294** (0.047) **< 0.001**	**0.130** (0.072) **0.072**	-	**−0.400** (0160) **0.013**	-	0.5467
Sy	**−0.290** (0.148) **0.050**	**0.607** (0.179) **0.001**	**−0.244** (0.069) **< 0.001**	**0.237** (0.057) **< 0.001**	0.097 (0.124) 0.435	0.144 (0.115) 0.212	−0.491 (0.320) 0.127	0.028 (0.209) 0.892	0.3571
Su	**−0.315** (0.105) **0.003**	**0.651** (0.144) **< 0.001**	**−0.183** (0.050) **< 0.001**	**0.327** (0.050) **< 0.001**	**0.146** (0.075) **0.052**	0.121 (0.080) 0.134	**−0.500** (0.167) **0.003**	-	0.511
Py	**−0.325** (0.115) **0.005**	**0.696** (0.159) **< 0.001**	**−0.268** (0.067) **< 0.001**	**0.248** (0.057) **< 0.001**	0.060 (0.107) 0.576	0.165 (0.106) 0.121	**−0.493** (0.184) **0.008**	-	0.3806
Pu	−0.124 (0.137) 0.366	**0.941** (0.159) **< 0.001**	**−0.151** (0.051) **0.003**	**0.243** (0.049) **< 0.001**	0.077 (0.085) 0.364	**0.234** (0.086) **0.007**	**−0.800** (0.287) **0.006**	**0.332** (0.143) **0.021**	0.5217
Ly	**−0.417** (0.111) **< 0.001**	**0.713** (0.153) **< 0.001**	**−0.295** (0.064) **< 0.001**	**0.234** (0.054) **< 0.001**	0.069 (0.103) 0.505	**0.187** (0.102) **0.069**	**−0.437** (0.177) **0.014**	-	0.4201
Lu	**−0.201** (0.104) **0.054**	**0.780** (0.142) **< 0.001**	**−0.183** (0.050) **< 0.001**	**0.266** (0.049) **< 0.001**	**0.180** (0.074) **0.015**	0.125 (0.079) 0.117	**−0.492** (0.165) **0.003**	-	0.521
Senergy	**−0.299** (0.140) **0.034**	**0.593** (0.170) **0.001**	**−0.168** (0.066) **0.011**	**0.288** (0.054) **< 0.001**	0.203 (0.118) 0.085	0.121 (0.110) 0.274	−0.360 (0.305) 0.239	−0.053 (0.199) 0.789	0.4343
Fenergy	-	0.112 (0.160) 0.485	-	**0.111** (0.064) **0.082**	**0.247** (0.117) **0.035**	-	-	-	0.109
Penergy	−0.089 (0.111) 0.424	0.047 (0.181) 0.797	0.075 (0.082) 0.359	−0.007 (0.069) 0.918	**0.274** (0.131) **0.037**	0.205 (0.130) 0.116	-	-	0.0677

Dependent variable | **Regression coefficients for standardized variables (Standard error) p-value** | | | | | | | **R^2**

	FX	Sex	Age	Weight	Height	Fx:Age	Fx:Sex	Fx:Height	
Lenergy	-	0.190 (0.178) 0.286	−0.096 (0.066) 0.144	**0.121 (0.067) 0.072**	0.157 (0.128) 0.221	-	-	-	0.0991

The bold-faced value means the variable is significant (p-value < 0.1). "−" indicates that the variable was not included in the model because the p-values would have been greater than 0.1

Table 3

The average AUCs of the classification models using stratified leave-one-group-out cross-validation for the fracture status prediction using $aBMD_{CT}$ and covariates, PC1 plus $aBMD_{CT}$ and covariates, and the nine FE parameters combined plus $aBMD_{CT}$ and covariates, respectively. All abbreviations of models are detailed in the Methods section

Model	Sample	Average AUC (Standard error)		
		$aBMDC_T$ + Covariates	PC1 + $aBMD_{CT}$ + Covariates	FE parameters combined + $aBMD_{CT}$ + Covariates
Logistic	Whole	0.699 (0.052)	**0.738 (0.043)**	0.724 (0.041)
	Male	0.727 (0.093)	**0.777 (0.093)**	0.745 (0.094)
	Female	0.608 (0.115)	0.623 (0.095)	**0.669 (0.089)**
LDA	Whole	0.696 (0.053)	**0.725 (0.046)**	0.708 (0.044)
	Male	0.727 (0.090)	**0.758 (0.099)**	0.751 (0.105)
	Female	0.608 (0.116)	0.615 (0.110)	**0.659 (0.080)**
PLS	Whole	0.700 (0.049)	**0.737 (0.045)**	0.731 (0.049)
	Male	0.719 (0.093)	**0.788 (0.113)**	0.753 (0.117)
	Female	0.617 (0.068)	**0.638 (0.087)**	0.611 (0.086)
RF	Whole	0.640 (0.045)	0.684 (0.042)	**0.727 (0.049)**
	Male	0.714 (0.107)	**0.745 (0.086)**	0.742 (0.072)
	Female	0.555 (0.060)	**0.577 (0.075)**	0.569 (0.075)
QDA	Whole	0.637 (0.073)	**0.695 (0.056)**	0.681 (0.052)
	Male	0.690 (0.110)	**0.707 (0.141)**	0.674 (0.101)
	Female	0.584 (0.085)	0.622 (0.067)	**0.649 (0.083)**
MDA	Whole	0.668 (0.056)	**0.702 (0.047)**	0.661 (0.048)
	Male	0.724 (0.078)	**0.719 (0.105)**	0.664 (0.102)
	Female	0.585 (0.074)	0.632 (0.066)	**0.645 (0.064)**
NNET	Whole	0.693 (0.061)	**0.731 (0.050)**	0.719 (0.046)
	Male	0.699 (0.109)	**0.744 (0.110)**	0.714 (0.093)
	Female	0.632 (0.115)	0.563 (0.136)	**0.622 (0.099)**
MARS	Whole	0.668 (0.052)	**0.717 (0.039)**	0.689 (0.054)
	Male	0.573 (0.127)	**0.724 (0.115)**	0.664 (0.137)
	Female	0.585 (0.086)	0.558 (0.094)	**0.667 (0.085)**

Model	Sample	Average AUC (Standard error)		
		$aBMDC_T$ + Covariates	PC1 + $aBMD_{CT}$ + Covariates	FE parameters combined + $aBMD_{CT}$ + Covariates
KNN	Whole	0.607 (0.120)	0.689 (0.074)	**0.691 (0.065)**
	Male	0.671 (0.128)	**0.744 (0.096)**	0.681 (0.128)
	Female	0.511 (0.078)	0.517 (0.109)	**0.554 (0.104)**

The bold-faced value means the largest AUCs among three comparison models