Title
Hip Fracture Discrimination Based on Statistical Multi-parametric Modeling (SMPM)

Journal
Annals of Biomedical Engineering, 47(11)

ISSN
0145-3068

Authors
Carballido-Gamio, Julio
Yu, Aihong
Wang, Ling
et al.

Publication Date
2019-11-01

DOI
10.1007/s10439-019-02298-x

Peer reviewed

HHS Public Access

Author manuscript

Ann Biomed Eng. Author manuscript; available in PMC 2020 November 01.

Published in final edited form as:

Ann Biomed Eng. 2019 November ; 47(11): 2199–2212. doi:10.1007/s10439-019-02298-x.

Hip Fracture Discrimination Based on Statistical Multi-Parametric Modeling (SMPM)

Julio Carballido-Gamio[1], **Aihong Yu**[2], **Ling Wang**[2], **Yongbin Su**[2], **Andrew J. Burghardt**[3], **Thomas F. Lang**[3], **Xiaoguang Cheng**[2]

[1]Department of Radiology, University of Colorado Anschutz Medical Campus, Aurora, CO, USA

[2]Department of Radiology, Beijing Jishuitan Hospital, Beijing, China

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

Abstract

Studies using quantitative computed tomography (QCT) and data-driven image analysis techniques have shown that trabecular and cortical volumetric bone mineral density (vBMD) can improve the hip fracture prediction of dual-energy X-ray absorptiometry areal BMD (aBMD). Here, we hypothesize that 1) QCT imaging features of shape, density and structure derived from data-driven image analysis techniques can improve the hip fracture discrimination of classification models based on mean femoral neck aBMD (Neck.aBMD), and 2) that data-driven cortical bone thickness (Ct.Th) features can improve the hip fracture discrimination of vBMD models. We tested our hypotheses using statistical multi-parametric modeling (SMPM) in a QCT study of acute hip fracture of 50 controls and 93 fragility fracture cases. SMPM was used to extract features of shape, vBMD, Ct.Th, cortical vBMD, and vBMD in a layer adjacent to the endosteal surface to develop hip fracture classification models with machine learning logistic LASSO. The performance of these classification models was evaluated in two aspects: 1) their hip fracture classification capability without Neck.aBMD, and 2) their capability to improve the hip fracture classification of the Neck.aBMD model. Assessments were done with 10-fold cross-validation, areas under the receiver operating characteristic curve (AUCs), differences of AUCs, and the integrated discrimination improvement (IDI) index. All LASSO models including SMPM-vBMD features, and the majority of models including SMPM-Ct.Th features performed significantly better than the Neck.aBMD model; and all SMPM features significantly improved the hip fracture discrimination of the Neck.aBMD model (Hypothesis 1). An interesting finding was that SMPM-features of vBMD also captured Ct.Th patterns, potentially explaining the superior classification performance of models based on SMPM-vBMD features (Hypothesis 2). Age, height and weight had a small impact on model performances, and the model of shape, vBMD and Ct.Th consistently yielded better performances than the Neck.aBMD models. Results of this study clearly support the relevance of bone density and quality on the assessment of hip fracture, and demonstrate their potential on patient and healthcare cost benefits.

Corresponding Author: Julio Carballido-Gamio, Department of Radiology, University of Colorado Anschutz Medical Campus, 12700 E 19th Ave, Mail Stop C278, Room 1208, Aurora CO 80045, Julio.Carballido-Gamio@ucdenver.edu, Tel: 303-724-6153.

Keywords

Quantitative computed tomography (QCT); hip; fracture; statistical multi-parametric modeling (SMPM); bone mineral density (BMD); cortical bone thickness (Ct.Th)

1. Introduction

Hip fractures have a devastating impact on the lives of elderly people and are imposing an increasing financial burden on a rapidly aging society.[9] In the United States, the proportion of the population older than 65 years will increase from 12.7% in 2000 to 20.3% in 2050,[47] and the number of bone fractures is expected to exceed 3 million by 2025 with associated costs in the order of $25.3 billion per year.[9] The individual risk of hip fracture depends on the structural integrity of the proximal femur and the likelihood of experiencing forces that exceed the bone strength.[31] With aging, the structural integrity of the proximal femur is compromised and the likelihood of falling is higher, predisposing an older person to an increase risk of hip fracture. Thus, the ability to identify individuals at high risk for fracture has become an imperative goal. Areal bone mineral density (aBMD) derived from dual-energy X-ray absorptiometry (DXA) is the clinical standard to assess osteoporosis status. Studies have consistently demonstrated that aBMD discriminatory and predictive capabilities of hip fracture are moderate based on area under the receiver operating characteristic curve (AUC) analyses (~0.80).[17, 5, 34, 2, 7, 4, 36, 8, 48, 18] The moderate performance of aBMD as a fracture predictor has been attributed to a function of factors it cannot measure, including microscopic factors such as bone microstructure, material properties and composition, and macroarchitectural factors such as bone shape, and the three-dimensional (3D) distribution of BMD and cortical bone thickness (Ct.Th). Furthermore, roughly half of the people who fracture do so with non-osteoporotic levels of aBMD.

Quantitative computed tomography (QCT) has become a leading imaging technique for osteoporosis-related research because it enables: 1) 3D macroarchitectural bone assessments in the two most clinically relevant skeletal sites, namely the proximal femur and the spine; 2) measurement of volumetric bone mineral density (vBMD); 3) distinction between cortical and trabecular bone compartments; 4) quantification of bone geometry and shape; 5) quantification of cortical bone thickness (Ct.Th); and 6) estimations of bone strength using subject-specific finite element modeling (FEM). All these characteristics therefore make QCT a unique medical imaging technique for comprehensive bone assessments in studies of etiology,[22] treatment, [32, 3, 42, 29, 35, 19, 21, 23, 20] and intervention.[33, 1] However, Black et al.[5] compared the hip fracture predictive capabilities of QCT with DXA in a large prospective study –the Osteoporotic Fractures in Men Study (MrOS)[6, 38]– finding that although specific QCT parameters were predictive of incident hip fracture independent of aBMD, adding QCT parameters to aBMD-based models did not improve hip fracture prediction. A potential explanation for this finding is that QCT parameters were computed based on predefined volumes of interest, specifically the femoral neck and a region including both the femoral neck and the trochanter. However, recent QCT publications based on Computational Anatomy techniques[13] have identified specific sub-regions in the proximal femur where

bone parameters have been associated with hip fracture.[37, 39, 11, 12, 8, 44] Computational Anatomy techniques are data-driven and therefore do not require predefinition of volumes of interest.

We hypothesized that QCT imaging features derived from data-driven image analysis techniques can improve the hip fracture discrimination of classification models based on mean femoral neck aBMD (Hypothesis 1). We also hypothesized that data-driven Ct.Th features can improve the hip fracture discrimination of vBMD models (Hypothesis 2). Here, we tested our hypotheses in a QCT study of acute hip fracture in Chinese women using statistical multi-parametric modeling (SMPM)[14] with parameters of shape, density and structure.[10] SMPM is an image analysis technique based on principal component analysis (PCA) of spatially normalized parametric maps. SMPM is data-driven and does not require predefinition of volumes of interest. SMPM enables the simultaneous analysis of shape with multiple parametric maps, basically extending statistical models of appearance[15] to multi-parametric intensity patches. Because of the inherent nature of PCA, SMPM with machine learning enable the identification of uncorrelated features that can be efficiently used for classification and prediction in studies of etiology, treatment, and intervention. Another key characteristic of SMPM is that it enables the display of uncorrelated parametric maps with concurrent features of shape.

2. Materials and Methods

2.1 Subjects

One hundred forty-three women were included in this study, 93 with and 50 without hip fracture. Inclusion and exclusion criteria were similar to those described by Li and colleagues.[37] Briefly, women with hip fracture were recruited from the emergency room, Department of Traumatology and Orthopedic Surgery, Beijing Jishuitan Hospital. Women were only included if: 1) their fracture had occurred within the last 48 hours (in order to minimize changes in vBMD and body composition factors due to the fracture); 2) were previously fully ambulatory, community-dwelling adults; and 3) fractures resulted from a low-energy fall from standing or sitting height. Controls were recruited from the surrounding community. Only women without conditions affecting bone metabolism, previous osteoporosis treatment, and without previous fractures were included in this study. Informed consent was obtained from all the participants after explaining them the nature of the study, and the study was reviewed and approved by the Internal Review Boards of the participating institutions.

2.2 Imaging

QCT imaging was performed with a Toshiba CT scanner (Toshiba Medical Systems Division, Tokyo, Japan). Bilateral hip acquisitions for all subjects were obtained in supine position from the top of the acetabulum to 3 cm below the lesser trochanter. At the time of scanning, a calibration phantom (Mindways Software, Austin, TX, USA) was placed under the CT table in order to convert Hounsfield units to equivalent concentrations of liquid K_2HPO_4 in mg/cm^3. Acquisitions were obtained with 120kVp, 125mAs, 50cm field of view,

and 512×512 matrix in spiral and standard reconstructions with an average of 226 slices per subject. Scans were reconstructed to voxels sizes of 0.781mm × 0.781mm × 1mm.

2.3 Image Processing

The left hip for controls and the contralateral unfractured hip for subjects with hip fracture were analyzed using an automatic pipeline.[10] The pipeline segmented the proximal femora, mirrored right hips, represented the periosteal boundaries with triangulated surfaces, and identified common anatomical volumes and surfaces of interest (femoral head, femoral neck, trochanteric region, and shaft). The pipeline therefore enabled the generation of multi-parametric maps encoding: 1) voxel-based vBMD, 2) surface-based Ct.Th, 3) surface-based mean cortical vBMD (Ct.vBMD), and 4) surface-based mean trabecular vBMD in a layer adjacent to the endosteal surface (EndoTb.vBMD).[10] Segmentations were manually corrected if necessary based on visual assessments, and the femoral head was excluded from the surface-based computations due to its thin cortical bone. Using the segmented vBMD maps and identified volumes of interest, QCT-derived aBMD maps were generated to compute mean femoral neck aBMD (Neck.aBMD) and mean total femur aBMD (Total.aBMD) values. Mean integral (cortical and trabecular) vBMD values for the femoral neck (Neck.vBMD) and total femur (Total.vBMD) were also computed.

SMPM is basically a statistical model of appearance where the intensity data can be multi-parametric. According to Cootes et al.,[15] a statistical model of appearance (shape and intensity) can be generated in three steps. First, PCA is performed on a set of aligned coordinates representing corresponding anatomical landmarks in the shapes of the population, yielding a statistical shape model (SSM).[16] Second, PCA is performed on a set of spatially normalized gray-level patches from the images of the population, yielding a statistical intensity model. And third, because there may be correlations between shape and intensity variations, an additional PCA is performed on the PCA scores of shape and intensity, yielding the statistical model of appearance. In order to include shape, SMPM follows these three steps, but in the second step either a single or multiple spatially normalized parametric maps are included, thus requiring an additional preprocessing step to deal with differences in units and dynamic ranges between the parameters.

To create the SSM,[16] corresponding anatomical landmarks were automatically identified on the periosteal surface of each segmented bone based on 3D registrations with a common template of the proximal femur.[11] Registrations were based on the segmented femora and included affine (3 translations, 3 rotations, and 3 scalings) and nonlinear transformations.[11] Using the identified anatomical landmarks, the periosteal surfaces of the population were concurrently aligned to remove differences in position and rotation, thus preserving size and shape differences. The SSM was then generated by applying PCA to the set of matched and spatially normalized vectors containing the coordinates of the periosteal surfaces in the population. Therefore, the SSM was represented by a mean femoral shape $\bar{\mathbf{x}}$ (column vector with coordinates of the periosteal surface of the mean shape), a set of orthogonal modes of shape variation $\mathbf{P}_{sh}$ (matrix where each column represents an orthogonal mode of shape variation), and a set of eigenvalues λ_{sh}. Using the mean shape and the orthogonal modes of

shape variation, the proximal femoral shape $\mathbf{x}_{S_n}$ of a specific subject was projected into the eigenspace of the SSM using Equation 1:

$$\mathbf{b}_{sh,S_n} = \mathbf{P}_{sh}^{\mathbf{T}}(\mathbf{x}_{S_n} - \bar{\mathbf{x}}) \quad (1)$$

In Equation 1, $\mathbf{b}_{sh,S_n}$ is a column vector representing the scores of the specific shape for each orthogonal mode of shape variation, i.e. the coordinates of the specific shape in the eigenspace of the SSM. Vectors like $\mathbf{b}_{sh,S_n}$ can be used for shape classification and prediction problems, so we evaluated their performance for hip fracture discrimination.

To create the multi-parametric feature model –second step towards SMPM that includes shape– the vBMD maps and the surface-based feature maps (Ct.Th, Ct.vBMD, and EndoTb.vBMD) were spatially normalized. Spatial normalization was performed with the transformations computed for the identification of the common anatomical landmarks. In order to adjust for differences in units and dynamic ranges between the features, each feature was normalized to zero mean and unit variance enabling their individual usage or their concatenation in a feature vector for PCA. PCA was then applied to the spatially- and scale-normalized feature vectors from the population yielding a mean feature vector $\bar{\mathbf{y}}$, a set of orthogonal modes of feature variation $\mathbf{P}_f$, and a set of eigenvalues λ_f. Using the mean feature vector and the set of orthogonal modes of feature variation, the feature vector $\mathbf{y}_{S_n}$ of a specific subject was projected into the eigenspace of the multi-parametric feature model using Equation 2:

$$\mathbf{b}_{f,S_n} = \mathbf{P}_f^{\mathbf{T}}(\mathbf{y}_{S_n} - \bar{\mathbf{y}}) \quad (2)$$

In Equation 2, $\mathbf{b}_{f,S_n}$ is a column vector representing the scores of the specific feature vector for each orthogonal mode of feature variation, i.e. the coordinates of the specific feature vector in the eigenspace of the multi-parametric feature model. Vectors like $\mathbf{b}_{f,S_n}$ can be used for feature classification and prediction problems, therefore, we evaluated their performance for hip fracture discrimination.

In the final step of SMPM incorporating shape, potential correlations between shape and feature variations were removed by first concatenating the individual shape and feature scores as:

$$\mathbf{b}_m = \begin{pmatrix} \mathbf{W}_{sh}\mathbf{b}_{sh} \\ \mathbf{b}_f \end{pmatrix} \quad (3)$$

and then by applying an additional PCA to these vectors. In Equation 3, $\mathbf{W}_{sh}$ is a diagonal matrix of weights for each shape parameter adjusting for the differences in units between the shape and feature models. This final step therefore also yields a mean shape-feature vector $\bar{\mathbf{z}}$, a set of orthogonal modes of shape-feature variation $\mathbf{P}_{sh-f}$, and a set of eigenvalues λ_{sh-f}. Using the mean shape-feature vector and the set of orthogonal modes of shape-feature

variation, the shape-feature vector $\mathbf{z}_{S_n}$ of a specific subject was projected into the eigenspace of the shape-feature model using Equation 4:

$$\mathbf{b}_{sh-f,S_n} = \mathbf{P}^{\mathbf{T}}_{sh-f}(\mathbf{z}_{S_n} - \bar{\mathbf{z}}) \quad (4)$$

In Equation 4, $\mathbf{b}_{sh-f,S_n}$ is a column vector representing the scores of the specific shape-feature vector for each orthogonal mode of shape-feature variation, i.e. the coordinates of the specific shape-feature vector in the eigenspace of the shape-multi-parametric feature model. Vectors like $\mathbf{b}_{sh-f,S_n}$ were also used in this work for hip fracture discrimination.

The relevance of performing an additional PCA step to generate the specific shape-feature scores as in Equation 4, is summarized in the Appendix.

2.4 Statistical Analysis

For each subject, vectors like those derived from Equations 1, 2 and 4 —$\mathbf{b}_{sh}$, $\mathbf{b}_{f}$, and $\mathbf{b}_{sh-f}$, which we will refer in general as **b**— have a total of *n-1* scores, where *n* is the number of subjects used for PCA. However, only few of those scores carry relevant fracture information. In order to select those relevant scores, the eigenvalues λ_{sh}, λ_{f}, and λ_{sh-f} were used to identify the contribution of each score towards the variance of the population, and select those representing 90% of the total variance —$\mathbf{b}_{sh,90\%}$, $\mathbf{b}_{f,90\%}$ and $\mathbf{b}_{sh-f,90\%}$, which we will refer in general as $\mathbf{b}_{90\%}$.

Scores in $\mathbf{b}_{90\%}$ vectors were used with machine learning logistic LASSO (Least Absolute Shrinkage and Selection Operator)[43] to perform automatic feature selection and hip fracture discrimination. Briefly, LASSO is a machine learning regression technique that shrinks the coefficients estimates towards zero, and actually forces some of the coefficient estimates to be exactly equal to zero, thus performing feature selection. Neck.aBMD, Total.aBMD, Neck.vBMD, and Total.vBMD are the most common clinical and research imaging parameters used to assess osteoporosis and hip fracture risk. Therefore, a LASSO model for each of these four parameters was evaluated with the aim of selecting a reference model for the SMPM models. Based on the performance of these models, the Neck.aBMD model was selected as the reference in this study. Therefore, scores from the $\mathbf{b}_{90\%}$ vectors, i.e. SMPM features, were evaluated in terms of their hip fracture discrimination capabilities without Neck.aBMD, and in terms of their improvement on hip fracture discrimination when added to the Neck.aBMD LASSO model. Furthermore, the hip fracture discrimination performance of SMPM features was also assessed after adding demographic and clinical variables —age, height and weight— to the different LASSO models. We used *Glmnet in Matlab*[40] for logistic LASSO (*elastic-net mixing parameter*=1, *response type*=binomial, and *penalty factors*=0 for Neck.aBMD, age, height and weight) and fracture discrimination was assessed with AUCs. Model comparisons were performed in terms of differences of AUCs and with the integrated discrimination improvement (IDI) index. AUCs were compared using the DeLong's nonparametric approach for two or more correlated AUCs. For AUCs and differences of AUCs, 95% confidence intervals were calculated; and for differences in AUCs

and IDI indices, p values were computed, which were considered significant at $p<0.05$. SMPM features included those of shape, vBMD, Ct.Th, Ct.vBMD, and EndoTb.vBMD.

Validation of the full approach to assess hip fracture discrimination was performed with stratified 10-fold cross-validation. Briefly, the population was randomly partitioned into 10 sub-groups with equal proportions of controls and fracture subjects. Then 10 iterations were executed. At each iteration, one sub-group represented the test data, and the remaining 9 sub-groups represented the training data. Using the training data, statistical multi-parametric models, **b** vectors and $\mathbf{b_{90\%}}$ vectors were generated, and LASSO was used for feature selection and generation of hip fracture discrimination models. Fracture discrimination was then assessed on the test sub-group. Results for all iterations were pooled to calculate the different performance metrics. By using cross-validation, we ensured that at each iteration, the model is tested on data that was not used for training. The objective of cross-validation is to avoid overfitting. Overfitting occurs when the models follow the errors, or noise, too closely, yielding a small training mean squared error (MSE) but a large test MSE, where MSE is defined by Equation 5:

$$MSE = \frac{1}{k}\sum_{i=1}^{k}\left(d_i - \hat{f}\,(q_i)\right)^2 \tag{5}$$

In Equation 5, k is the number of data points, d_i is the true value for the ith observation, and $\hat{f}(q_i)$ is the prediction that $\hat{f}$ gives for the ith observation.

The full pipeline was implemented in MATLAB (The MathWorks, Inc. Natick, MA) on an iMac computer with a 4GHz Intel Core i7 processor and 32GB of RAM, and it is briefly outlined in the diagram of Figure 1.

3. Results

Table 1 shows the subject characteristics for the women involved in this study. Controls and women with hip fracture had significantly different age ($p<0.001$) and height ($p=0.02$), but not weight.

Tables 2 and 3 show the hip fracture discrimination performance of different LASSO models based on Neck.aBMD, Total.aBMD, Neck.vBMD, Total.vBMD, and SMPM features without and with the inclusion of demographic and clinical variables (age, height, and weight), respectively. The performance of the LASSO models is given in terms of AUCs, and in terms of differences in AUCs and IDI indexes with respect to the Neck.aBMD models (without and with age + height + weight). According to Table 2, Total.aBMD, Neck.vBMD, Total.vBMD, Ct.vBMD, and shape + Ct.vBMD yielded similar performances than Neck.aBMD; while all LASSO models including SMPM-vBMD features performed significantly better. Table 2 also shows that with the exception of the shape + Ct.Th model, models including SMPM-Ct.Th features also performed significantly better than Neck.aBMD. The rest of the parameters in Table 2 yielded inconsistent results between AUC differences and IDI indexes. When age, height and weight were added to the LASSO models, only shape + vBMD + Ct.Th yielded significantly higher fracture discrimination

than the reference model (Hypothesis 2), while the rest of the parameters performed similar than the reference model or yielded inconsistent results between AUC differences and IDI indexes (Table 3).

Tables 4 and 5 summarize the improvements on hip fracture discrimination when SMPM features (i.e. principal component scores $\mathbf{b}_{90\%}$) were added to the LASSO models of Neck.aBMD, without and with demographic and clinical variables (age, height, and weight), respectively. As in Tables 2 and 3, the performances of the models are given in terms of AUCs, and in terms of differences in AUCs and IDI indexes with respect to the Neck.aBMD models. According to Table 4, when SMPM features were added to the Neck.aBMD model, significant hip fracture discrimination improvements were observed in all instances in terms of both AUCs (0.041–0.081; $p<0.001$–$p<0.05$) and IDI indexes (0.064–0.243; $p<0.001$). When age, height and weight were added to the Neck.aBMD model (Table 5), all models incorporating EndoTb.vBMD improved the performance of the Neck.aBMD + age + height + weight LASSO model. Other models showing improvement were: 1) Ct.vBMD, 2) shape + Ct.Th, and 3) shape + vBMD + Ct.Th. The rest of the parameters showed similar performance than the reference model or inconsistent results between differences in AUCs and IDI indexes.

Figures 2-4 show some representative SMPM features selected as relevant predictors by LASSO according to Tables 2-5. Figure 2 shows coronal cross-sections of the mean statistical model of vBMD and the mean minus and plus 2.5 standard deviations along the directions of the first and second principal components. These two modes were consistently selected as predictors in the different k folds. According to Tables 2 and 4, principal component scores of vBMD yielded LASSO models that performed better and that improved hip fracture discrimination with respect to Neck.aBMD, respectively. While Figure 2A shows clear vBMD differences on opposite directions of the first principal component effectively in all proximal femoral regions in both the cortical and the trabecular bone compartments, Figure 2B shows more subtle differences on opposite directions of the second principal component, mainly at the intertrochanteric region and the medial cortex. Interestingly, both modes seem to also capture cortical bone thickness patterns. In average, the control group showed negative scores (patterns of higher vBMD and thicker cortices), while the fracture group showed positive scores (patterns of lower vBMD and thinner cortices) for both principal components of vBMD.

In Figure 3, posterior and anterior views of the mean statistical model of Ct.Th together with the mean minus and plus 2.5 standard deviations in the direction of the first principal component are shown. Mode 1 was consistently selected as predictor in the different k folds. According to Tables 2 and 4, principal component scores of Ct.Th yielded LASSO models that performed better and that improved hip fracture discrimination with respect to Neck.aBMD, respectively. Figure 3 shows clear cortical bone thickness differences on opposite directions of the first principal component in focal regions, particularly in: 1) the medial cortex, 2) the femoral neck, and 3) and anteriorly in the greater trochanter. In average, the control group showed negative scores (patterns of thicker Ct.Th), while the fracture group showed positive scores (patterns of thinner Ct.Th) for the first principal component of Ct.Th.

In terms of shape, Figure 4 shows the mean minus and plus 2.5 standard deviations of the SSM along the directions of the first (Figure 4A) and second principal components (Figure 4B), which were consistently selected by LASSO. According to Table 4, shape features significantly improved hip fracture discrimination on top of Neck.aBMD. As expected, Figure 4A captures size differences in the first mode, while Figure 4B captures differences in the neck-shaft angle in the second mode. In average, controls showed negative shape scores (towards blue shapes), and fracture cases positive scores (towards gray shapes).

4. Discussion

SMPM is a data-driven image analysis technique that enables the incorporation of multiple spatially normalized parameters into models that generate uncorrelated scores suitable for feature selection and computation of classification and prediction models. In this work, we applied SMPM with machine learning logistic LASSO to discriminate hip fracture in a study of acute hip fracture in Chinese women, including 50 controls and 93 with hip fracture. The hip fracture discrimination performance of LASSO models based on SMPM features without QCT-derived Neck.aBMD, and when SMPM features were added to the Neck.aBMD model was evaluated using 10-fold cross-validation to prevent overfitting. SMPM features included those of shape, vBMD, Ct.Th, Ct.vBMD, EndoTb.vBMD, and some of their combinations. Performances were evaluated in terms of AUCs, and comparisons with the Neck.aBMD reference model were done in terms of differences of AUCs and IDI indexes. The effects of incorporating demographic and clinical variables (age, height, and weight) to these models were also evaluated.

The added value of data-driven techniques incorporating 2D shape and aBMD features in the assessment of hip fracture has been demonstrated in DXA studies in the form of statistical models of shape, active shape modeling and statistical models of appearance.[26, 25, 2, 24] The added value of data-driven techniques incorporating vBMD, 3D shape, cortical mass surface density and endocortical trabecular bone mineral density in the assessment of hip fracture has also been demonstrated in QCT studies, in this case, in the form of regions of interest derived from voxel-based morphometry (VBM),[37] statistical models of appearance,[41, 8] separate models of shape and vBMD,[46] and regions of interest derived from cortical bone mapping.[44] However, our study is the first QCT study to simultaneously evaluate data-driven features of shape, density and structure, and their combinations, for the assessment of hip fracture.

LASSO models based on SMPM features yielded high levels of discrimination without Neck.aBMD with AUCs ranging from 0.789 to 0.927 (Table 2). Results also showed significant hip fracture discrimination improvements when SMPM features of shape, density or structure were added to the LASSO model based on Neck.aBMD, with ΔAUCs ranging from 0.041 to 0.081, and IDI indices ranging from 0.064 to 0.243 (Table 4), validating Hypothesis 1. The incorporation of age, height and weight to the different LASSO models, yielded non-significant AUC improvements and minor IDI changes (comparison of Tables 2 and 3; ~IDI=0.05; ~$p<0.001$; data not shown). Furthermore, although SMPM features of Ct.Th increased neither the AUC nor the IDI of the vBMD and shape + vBMD models (with or without age + height + weight; data not shown), the shape + vBMD + Ct.Th + age +

height + weight SMPM model was the only model that performed significantly better than the Neck.aBMD + age + height + weight model (Table 3; ΔAUC=0.054 and IDI=0.082), partially validating Hypothesis 2. The addition of shape, vBMD and Ct.Th SMPM features to the Neck.aBMD + age + height + weight model also significantly improved hip fracture discrimination (Table 5; ΔAUC=0.043; IDI=0.159).

Our AUCs of 0.844 and 0.872 for Neck.aBMD and Neck.aBMD + age + height + weight, respectively, were in agreement with those in the literature.[8, 44] Bredbenner et al. reported an AUC=0.82 for hip fracture prediction in men based on total hip aBMD, and an AUC=0.83 when using total hip aBMD + age + BMI;[8] while Treece et al. reported an AUC=0.78 for hip fracture prediction in men based on aBMD + age + height + clinical site.[44] Our SMPM results were also in agreement with those of the few data-driven QCT studies of hip fracture in the literature.[37, 8] The results presented here for hip fracture discrimination using SMPM features of vBMD (AUC=0.915) were similar to those in the hip fracture discrimination study of Li et al. using VBM (AUC=0.92);[37] while the results for shape + vBMD (AUC=0.924) were also in agreement with those of the statistical model appearance of Bredbenner et al. (AUC=0.94).[8] However, it is important to note that there are two main differences between the statistical model of appearance of Bredbenner et al. and the SMPM presented here. First, in this work, we described the proximal femoral shapes only with landmarks on the femoral surfaces, while Bredbenner et al. described the proximal femoral shapes with volumetric meshes, i.e. with landmarks on the femoral surfaces and with landmarks within the femora. Second, in this work, the 10-fold cross-validation was done for the full image processing pipeline: SMPM, feature selection, computation of LASSO models, and testing; while Bredbenner et al. used 10-fold cross-validation only for feature selection, computation of prediction models, and testing. Nevertheless, it is important to highlight the similarity of the results in three different data-driven studies incorporating vBMD, which underscores the robustness of these image analysis techniques.

QCT imaging features of shape, density and structure derived from SMPM improved the hip fracture discrimination of classification models based on mean femoral neck aBMD (Neck.aBMD), validating Hypothesis 1. According to Burge et al.[9] the total number of female incident hip fractures in the US in 2005 was 222,753. Based on our Table 5, by adding SMPM features to a model based on Neck.aBMD + age + height + weight, additional 4.1% (EndoTb.vBMD)-5.4% (Shape+EndoTb.vBMD) hip fractures could be detected. These numbers represent 9,132-12,028 hip fractures which represent 382-503 US million dollars associated with hip fracture costs that could be saved in the ideal scenario that this methodology works similarly for prediction purposes. These numbers clearly reflect the practical importance of these findings. Looking at Figures 2-4, the principal modes that were automatically selected by LASSO captured both global and local patterns, suggesting that both global and local bone deterioration and adaptation with aging might play significant roles on the assessment of bone strength. In addition, Figure 2 also showed an interesting finding with the principal modes of vBMD capturing Ct.Th patterns. vBMD and Ct.Th. are perhaps the most representative surrogate measures of bone strength that can be derived from QCT. Since bone strength plays a critical role in bone fracture, Hypothesis 2 was based on the idea that by incorporating two surrogate measures of bone strength, vBMD and Ct.Th, the discrimination between controls and fracture cases should be higher than that

based on models relying on vBMD alone. While the vBMD + Ct.Th model did not perform better than the vBMD model, Ct.Th did contribute to a better classification performance of vBMD models based on Figure 2 and Table 3, partially validating Hypothesis 2. A practical consideration of these results is that when using SMPM of vBMD maps for classification purposes, the inclusion of Ct.Th maps might not be needed to improve the performance since this information is embedded in the SMPM-vBMD features. In addition, consistent with previous QCT studies of hip fracture, the femoral neck played a major role on the assessment of hip fracture.[28] Results also demonstrated that processing the same vBMD data with a different method —with SMPM— led to a significant improvement on hip fracture discrimination compared to Neck.vBMD and Total.vBMD.

This study has four main limitations. First, although women with hip fracture were scanned within 48 hours of the occurrence of hip fracture, this is not a prospective study of incident hip fracture, and due to the lack of a baseline scan before the occurrence of fractures, results cannot be interpreted in the context of hip fracture prediction. Future studies are therefore needed to evaluate SMPM for hip fracture prediction in prospective studies such as the AGES-Reykjavik[27] or MrOS.[6, 38] The second limitation of this study is that the population was limited to women, and results cannot be extrapolated to men due to known sex differences in bone strength.[30] Therefore, future studies should evaluate SMPM in male populations such as that of the MrOS study.[6, 38] The third limitation is related to the ethnicity. All women involved in this study were Chinese, however, there are also known differences in bone characteristics between Asian and Caucasian women.[45] Therefore, the performance of SMPM should also be evaluated in populations of different ethnicities. And the fourth limitation is the significant difference in age between the control and fracture groups. Although we tried to explore the effect of age in our classification models and it was found to be small (Tables 3 and 5), the significant age difference might reduce the impact of our findings.

In conclusion, we have presented the application of SMPM for the assessment of hip fracture discrimination in a study of acute hip fracture in Chinese women. SMPM enables the incorporation of multiple spatially normalized parameters and shape into models that use PCA and feature selection making them suitable for classification and prediction problems. In this study, SMPM features discriminated hip fracture without Neck.aBMD, improved hip fracture discrimination based on Neck.aBMD (Hypothesis 1), and demonstrated the relevance of Ct.Th on classification models based on vBMD (Hypothesis 2). The findings of this study clearly support the relevance of bone density and quality on the assessment of hip fracture, and demonstrate their potential on patient and healthcare cost benefits.

Acknowledgements

This work was supported by the NIH/NIAMS under grants R01AR068456 and R01AR064140. This study was also supported by grants from the National Natural Science Foundation of China (81071131), the Beijing Bureau of Health 215 Program (2013-3-033; 2009-2-03), Beijing Technology Foundation for Selected Overseas Chinese Scholar and Beijing Talents Fund (2015000021467), Capital Characteristic Clinic Project (Z141107002514072).

This is a post-peer-review, pre-copyedit version of an article published in Annals of Biomedical Engineering. The final authenticated version is available online at: http://dx.doi.org/10.1007/s10439-019-02298-x.

Appendix

This section has the purpose of supporting the necessity of a third PCA step to address potential correlations between shape and feature principal component scores.

In this study, for the first 10 principal components, out of 100 correlations of principal component scores: 1) 37% between shape and vBMD, 2) 43% between shape and Ct.Th, 3) 29% between shape and Ct.vBMD, and 4) 40% between shape and EndoTb.vBMD, were significant.

References

1. Allison SJ, Poole KES, Treece GM, Gee AH, Tonkin C, Rennie WJ, Folland JP, Summers GD and Brooke-Wavell K. The Influence of High-Impact Exercise on Cortical and Trabecular Bone Mineral Content and 3D Distribution Across the Proximal Femur in Older Men: A Randomized Controlled Unilateral Intervention. Journal of Bone and Mineral Research 30(9):1709–1716, 2015. [PubMed: 25753495]
2. Baker-LePain JC, Luker KR, Lynch JA, Parimi N, Nevitt MC and Lane NE. Active Shape Modeling of the Hip in the Prediction of Incident Hip Fracture. Journal of Bone and Mineral Research 26(3): 468–474, 2011. [PubMed: 20878772]
3. Bauer DC, Garnero P, Bilezikian JP, Greenspan SL, Ensrud KE, Rosen CJ, Palermo L and Black DM. Short-term changes in bone turnover markers and bone mineral density response to parathyroid hormone in postmenopausal women with osteoporosis. J Clin Endocrinol Metab 91(4):1370–1375, 2006. [PubMed: 16449339]
4. Berry SD, Samelson EJ, Pencina MJ, McLean RR, Cupples LA, Broe KE and Kiel DP. Repeat bone mineral density screening and prediction of hip and major osteoporotic fracture. JAMA 310(12): 1256–1262, 2013. [PubMed: 24065012]
5. Black DM, Bouxsein ML, Marshall LM, Cummings SR, Lang TF, Cauley JA, Ensrud KE, Nielson CM, Orwoll ES and Osteoporotic G Fractures in Men Research. Proximal femoral structure and the prediction of hip fracture in men: a large prospective study using QCT. J Bone Miner Res 23(8): 1326–1333, 2008. [PubMed: 18348697]
6. Blank JB, Cawthon PM, Carrion-Petersen ML, Harper L, Johnson JP, Mitson E and Delay RR. Overview of recruitment for the osteoporotic fractures in men study (MrOS). Contemp Clin Trials 26(5):557–568, 2005. [PubMed: 16085466]
7. Bousson VD, Adams J, Engelke K, Aout M, Cohen-Solal M, Bergot C, Haguenauer D, Goldberg D, Champion K, Aksouh R, Vicaut E and Laredo JD. In Vivo Discrimination of Hip Fracture With Quantitative Computed Tomography: Results From the Prospective European Femur Fracture Study (EFFECT). Journal of Bone and Mineral Research 26(4):881–893, 2011. [PubMed: 20939025]
8. Bredbenner TL, Mason RL, Havill LM, Orwoll ES, Nicolella DP and Osteoporotic S Fractures in Men. Fracture risk predictions based on statistical shape and density modeling of the proximal femur. J Bone Miner Res 29(9):2090–2100, 2014. [PubMed: 24692132]
9. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A and Tosteson A. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005-2025. J Bone Miner Res 22(3):465–475, 2007. [PubMed: 17144789]
10. Carballido-Gamio J, Bonaretti S, Saeed I, Harnish R, Recker R, Burghardt AJ, Keyak JH, Harris T, Khosla S and Lang TF. Automatic multi-parametric quantification of the proximal femur with quantitative computed tomography. Quant Imaging Med Surg 5(4):552–568, 2015. [PubMed: 26435919]
11. Carballido-Gamio J, Harnish R, Saeed I, Streeper T, Sigurdsson S, Amin S, Atkinson EJ, Therneau TM, Siggeirsdottir K, Cheng X, Melton LJ 3rd, Keyak J, Gudnason V, Khosla S, Harris TB and Lang TF. Proximal femoral density distribution and structure in relation to age and hip fracture risk in women. J Bone Miner Res 28(3):537–546, 2013. [PubMed: 23109068]

12. Carballido-Gamio J, Harnish R, Saeed I, Streeper T, Sigurdsson S, Amin S, Atkinson EJ, Therneau TM, Siggeirsdottir K, Cheng X, Melton LJ 3rd, Keyak JH, Gudnason V, Khosla S, Harris TB and Lang TF. Structural patterns of the proximal femur in relation to age and hip fracture risk in women. Bone 57(1):290–299, 2013. [PubMed: 23981658]
13. Carballido-Gamio J and Nicolella DP. Computational anatomy in the study of bone structure. Curr Osteoporos Rep 11(3):237–245, 2013. [PubMed: 23722733]
14. Carballido-Gamio J, Yu A, Wang L, Yongbin S, Lang TF and Cheng X. Fracture Risk Estimation with Statistical Multi-Parametric Modeling. ASBMR Annual Meeting, 2016.
15. Cootes TF and Taylor CJ. Statistical models of appearance for medical image analysis and computer vision. Medical Imaging: 2001: Image Processing, Pts 1-3 2(27):236–248, 2001.
16. Cootes TF, Taylor CJ, Cooper DH and Graham J. Active Shape Models - Their Training and Application. Computer Vision and Image Understanding 61(1):38–59, 1995.
17. Crabtree NJ, Kroger H, Martin A, Pols HA, Lorenc R, Nijs J, Stepan JJ, Falch JA, Miazgowski T, Grazio S, Raptou P, Adams J, Collings A, Khaw KT, Rushton N, Lunt M, Dixon AK and Reeve J. Improving risk assessment: hip geometry, bone mineral distribution and bone strength in hip fracture cases and controls. The EPOS study. European Prospective Osteoporosis Study. Osteoporos Int 13(1):48–54, 2002. [PubMed: 11883408]
18. Dong XLN, Pinninti R, Lowe T, Cussen P, Ballard JE, Di Paolo D and Shirvaikar M. Random field assessment of inhomogeneous bone mineral density from DXA scans can enhance the differentiation between postmenopausal women with and without hip fractures. Journal of Biomechanics 48(6):1043–1051, 2015. [PubMed: 25683520]
19. Eastell R, Lang T, Boonen S, Cummings S, Delmas PD, Cauley JA, Horowitz Z, Kerzberg E, Bianchi G, Kendler D, Leung P, Man Z, Mesenbrink P, Eriksen EF, Black DM and Trial HPF. Effect of once-yearly zoledronic acid on the spine and hip as measured by quantitative computed tomography: results of the HORIZON Pivotal Fracture Trial. Osteoporos Int 21(7):1277–1285, 2010. [PubMed: 19802508]
20. Engelke K, Fuerst T, Dardzinski B, Kornak J, Ather S, Genant HK and de Papp A. Odanacatib treatment affects trabecular and cortical bone in the femur of postmenopausal women: results of a two-year placebo-controlled trial. J Bone Miner Res 30(1):30–38, 2015. [PubMed: 24898537]
21. Engelke K, Fuerst T, Dasic G, Davies RY and Genant HK. Regional distribution of spine and hip QCT BMD responses after one year of once-monthly ibandronate in postmenopausal osteoporosis. Bone 46(6):1626–1632, 2010. [PubMed: 20226286]
22. Engelke K, Lang T, Khosla S, Qin L, Zysset P, Leslie WD, Shepherd JA and Schousboe JT. Clinical Use of Quantitative Computed Tomography (QCT) of the Hip in the Management of Osteoporosis in Adults: the 2015 ISCD Official Positions-Part I. Journal of Clinical Densitometry 18(3):338–358, 2015. [PubMed: 26277851]
23. Genant HK, Libanati C, Engelke K, Zanchetta JR, Hoiseth A, Yuen CK, Stonkus S, Bolognese MA, Franek E, Fuerst T, Radcliffe HS and McClung MR. Improvements in hip trabecular, subcortical, and cortical density and mass in postmenopausal women with osteoporosis treated with denosumab. Bone 56(2):482–488, 2013. [PubMed: 23871852]
24. Goodyear SR, Barr RJ, McCloskey E, Alesci S, Aspden RM, Reid DM and Gregory JS. Can we improve the prediction of hip fracture by assessing bone structure using shape and appearance modelling? Bone 53(1):188–193, 2013. [PubMed: 23220597]
25. Gregory JS, Stewart A, Undrill PE, Reid DM and Aspden RM. Bone shape, structure, and density as determinants of osteoporotic hip fracture: a pilot study investigating the combination of risk factors. Invest Radiol 40(9):591–597, 2005. [PubMed: 16118552]
26. Gregory JS, Testi D, Stewart A, Undrill PE, Reid DM and Aspden RM. A method for assessment of the shape of the proximal femur and its relationship to osteoporotic hip fracture. Osteoporos Int 15(1):5–11, 2004. [PubMed: 14605797]
27. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF, Hoffman HJ and Gudnason V. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. Am J Epidemiol 165(9): 1076–1087, 2007. [PubMed: 17351290]

28. Johannesdottir F, Turmezei T and Poole KE. Cortical bone assessed with clinical computed tomography at the proximal femur. J Bone Miner Res 29(4):771–783, 2014. [PubMed: 24677244]
29. Keaveny TM, Hoffmann PF, Singh M, Palermo L, Bilezikian JP, Greenspan SL and Black DM. Femoral bone strength and its relation to cortical and trabecular changes after treatment with PTH, alendronate, and their combination as assessed by finite element analysis of quantitative CT scans. J Bone Miner Res 23(12):1974–1982, 2008. [PubMed: 18684084]
30. Keyak JH, Sigurdsson S, Karlsdottir G, Oskarsdottir D, Sigmarsdottir A, Zhao S, Kornak J, Harris TB, Sigurdsson G, Jonsson BY, Siggeirsdottir K, Eiriksdottir G, Gudnason V and Lang TF. Male-female differences in the association between incident hip fracture and proximal femoral strength: a finite element analysis study. Bone 48(6):1239–1245, 2011. [PubMed: 21419886]
31. Keyak JH, Sigurdsson S, Karlsdottir GS, Oskarsdottir D, Sigmarsdottir A, Kornak J, Harris TB, Sigurdsson G, Jonsson BY, Siggeirsdottir K, Eiriksdottir G, Gudnason V and Lang TF. Effect of finite element model loading condition on fracture risk assessment in men and women: the AGES-Reykjavik study. Bone 57(1):18–29, 2013. [PubMed: 23907032]
32. Lane NE, Sanchez S, Modin GW, Genant HK, Pierini E and Arnaud CD. Bone mass continues to increase at the hip after parathyroid hormone treatment is discontinued in glucocorticoid-induced osteoporosis: results of a randomized controlled clinical trial. J Bone Miner Res 15(5):944–951, 2000. [PubMed: 10804025]
33. Lang TF, Saeed IH, Streeper T, Carballido-Gamio J, Harnish RJ, Frassetto LA, Lee SM, Sibonga JD, Keyak JH, Spiering BA, Grodsinsky CM, Bloomberg JJ and Cavanagh PR. Spatial heterogeneity in the response of the proximal femur to two lower-body resistance exercise regimens. J Bone Miner Res 29(6):1337–1345, 2014. [PubMed: 24293094]
34. Leslie WD, Pahlavan PS, Tsang JF, Lix LM and Manitoba Bone Density P. Prediction of hip and other osteoporotic fractures from hip geometry in a large clinical cohort. Osteoporos Int 20(10): 1767–1774, 2009. [PubMed: 19238304]
35. Lewiecki EM, Keaveny TM, Kopperdahl DL, Genant HK, Engelke K, Fuerst T, Kivitz A, Davies RY and Fitzpatrick LA. Once-monthly oral ibandronate improves biomechanical determinants of bone strength in women with postmenopausal osteoporosis. J Clin Endocrinol Metab 94(1):171–180, 2009. [PubMed: 18840641]
36. Li GW, Chang SX, Xu Z, Chen Y, Bao H and Shi X. Prediction of hip osteoporotic fractures from composite indices of femoral neck strength. Skeletal Radiol 42(2):195–201, 2013. [PubMed: 22714125]
37. Li W, Kornak J, Harris T, Keyak J, Li C, Lu Y, Cheng X and Lang T. Identify fracture-critical regions inside the proximal femur using statistical parametric mapping. Bone 44(4):596–602, 2009. [PubMed: 19130910]
38. Orwoll E, Blank JB, Barrett-Connor E, Cauley J, Cummings S, Ensrud K, Lewis C, Cawthon PM, Marcus R, Marshall LM, McGowan J, Phipps K, Sherman S, Stefanick ML and Stone K. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study--a large observational study of the determinants of fracture in older men. Contemp Clin Trials 26(5):569–585, 2005. [PubMed: 16084776]
39. Poole KE, Treece GM, Mayhew PM, Vaculik J, Dungl P, Horak M, Stepan JJ and Gee AH. Cortical thickness mapping to identify focal osteoporosis in patients with hip fracture. PLoS One 7(6):e38466, 2012. [PubMed: 22701648]
40. Qian J, Hastie T, Friedman J, Tibshirani R and Simon N. “Glmnet for Matlab (2013).” from http://www.stanford.edu/~hastie/glmnet_matlab/.
41. Schuler B, Fritscher KD, Kuhn V, Eckstein F, Link TM and Schubert R. Assessment of the individual fracture risk of the proximal femur by using statistical appearance models. Med Phys 37(6):2560–2571, 2010. [PubMed: 20632568]
42. Sellmeyer DE, Black DM, Palermo L, Greenspan S, Ensrud K, Bilezikian J and Rosen CJ. Hetereogeneity in skeletal response to full-length parathyroid hormone in the treatment of osteoporosis. Osteoporos Int 18(7):973–979, 2007. [PubMed: 17333451]
43. Tibshirani R Regression shrinkage and selection via the lasso. J. Royal. Statist. Soc B 58(1):267–288, 1996.

44. Treece GM, Gee AH, Tonkin C, Ewing SK, Cawthon PM, Black DM, Poole KE and Osteoporotic S Fractures in Men. Predicting Hip Fracture Type With Cortical Bone Mapping (CBM) in the Osteoporotic Fractures in Men (MrOS) Study. J Bone Miner Res 30(11):2067–2077, 2015. [PubMed: 25982802]

45. Walker MD, Saeed I, McMahon DJ, Udesky J, Liu G, Lang T and Bilezikian JP. Volumetric bone mineral density at the spine and hip in Chinese American and White women. Osteoporos Int 23(10):2499–2506, 2012. [PubMed: 22147209]

46. Whitmarsh T, Fritscher KD, Humbert L, Del Rio Barquero LM, Roth T, Kammerlander C, Blauth M, Schubert R and Frangi AF. A statistical model of shape and bone mineral density distribution of the proximal femur for fracture risk assessment. Med Image Comput Comput Assist Interv 14(Pt 2):393–400, 2011. [PubMed: 21995053]

47. Wiener JM and Tilly J. Population ageing in the United States of America: implications for public programmes. Int J Epidemiol 31(4):776–781, 2002. [PubMed: 12177018]

48. Yang L, Udall WJM, McCloskey EV and Eastell R. Distribution of bone density and cortical thickness in the proximal femur and their association with hip fracture in postmenopausal women: a quantitative computed tomography study. Osteoporosis International 25(1):251–263, 2014. [PubMed: 23719860]

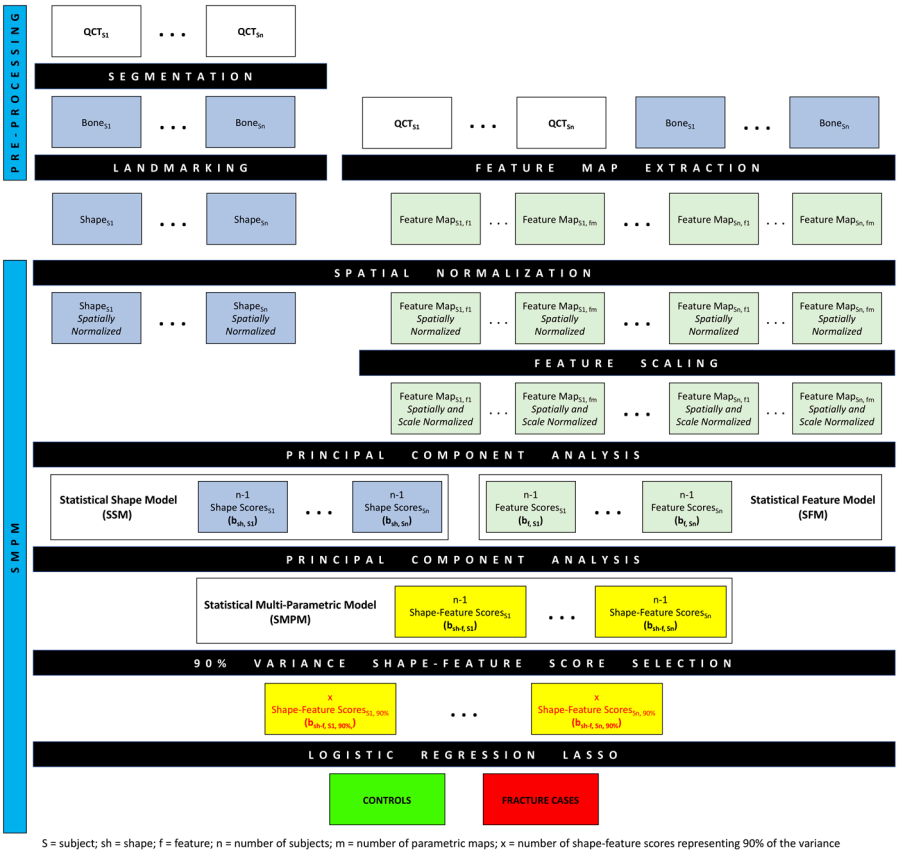

Figure 1.
Diagram outlining the steps of the SMPM pipeline implemented in this study.

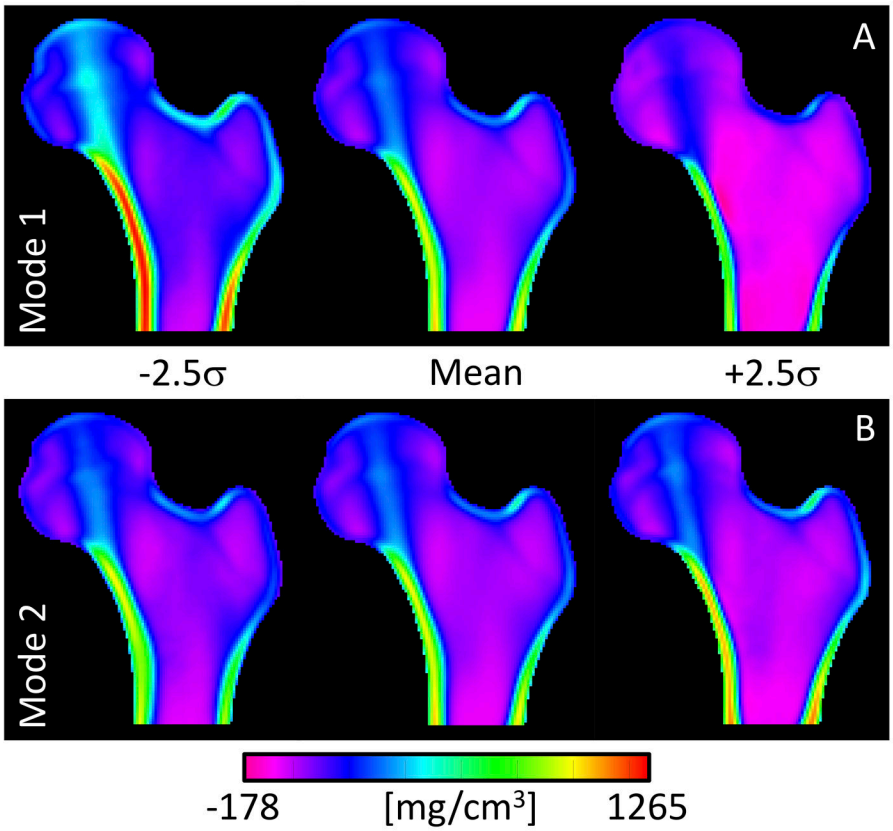

Figure 2.

Coronal cross-sections of the mean model of vBMD (center) and the mean minus (left) and plus (right) 2.5 standard deviations along the directions of the first (A) and second (B) principal components of the statistical model of vBMD. Interestingly, both modes also captured Ct.Th patterns.

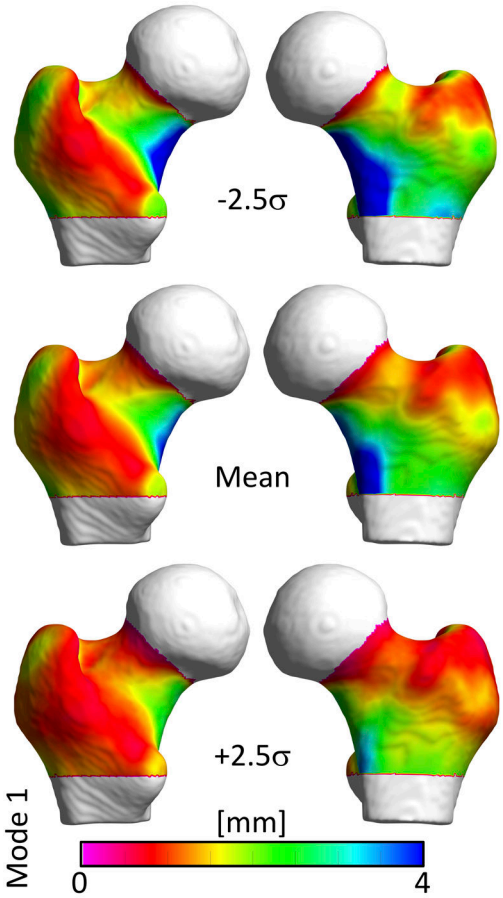

Figure 3.

Posterior and anterior proximal femoral views of the mean model of Ct.Th (middle row) and the mean minus (top) and plus (bottom) 2.5 standard deviations along the direction of the first principal component of the statistical model of Ct.Th.

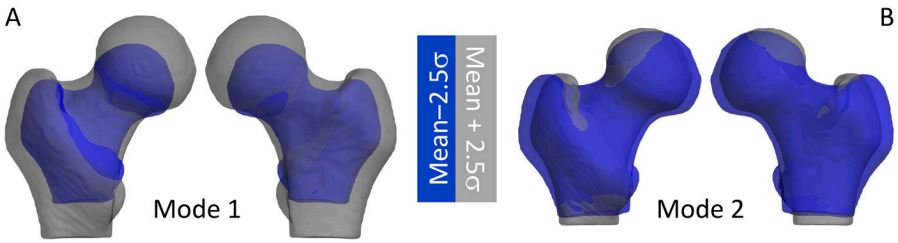

Figure 4.

Posterior and anterior proximal femoral views of the overlay of the mean shape minus 2.5 standard deviations (blue) and the mean shape plus 2.5 standard deviations (gray) along the first (A) and second (B) principal components of the SSM.

Table 1.

Descriptive Statistics for the Women Involved in this Study – Mean±SD

Group	Age [years]*	Height [cm]†	Weight [Kg]
Controls	63.7±7.0	156.8±6.0	63.4±9.1
Fracture	71.7±12.2	159.1±5.5	59.6±13.1

*Significantly different between groups: p<0.001

†Significantly different between groups: p=0.02

50 controls - 93 hip fracture cases

Table 2.

Hip Fracture Discrimination and Comparisons to Neck.aBMD

Initial Features in the LASSO Model	AUC [95% CI]	AUC Difference [95% CI] (p-Value)	IDI Index (p-Value)
Neck.aBMD	0.844 [0.780, 0.908]		
Total.aBMD	0.845 [0.781, 0.909]	0.001 [−0.030, 0.031] (NS)	−0.024 (NS)
Neck.vBMD	0.842 [0.778, 0.907]	−0.002 [−0.031, 0.027] (NS)	−0.021 (NS)
Total.vBMD	0.823 [0.755, 0.892]	0.0002 [−0.032, 0.033] (NS)	−0.020 (NS)
Shape	0.789 [0.714, 0.863]	−0.056 [−0.164, 0.053] (NS)	−0.270 (<0.001)
vBMD	0.915 [0.869, 0.962]	0.071 [0.026, 0.116] (<0.01)	0.083 (<0.01)
Ct.Th	0.915 [0.869, 0.962]	0.071 [0.009, 0.133] (<0.05)	0.100 (<0.05)
Ct.vBMD	0.884 [0.830, 0.939]	0.040 [−0.023, 0.104] (NS)	0.048 (NS)
EndoTb.vBMD	0.914 [0.867, 0.961]	0.070 [0.019, 0.121] (<0.01)	0.073 (0.053)
vBMD + CtTh	0.919 [0.874, 0.964]	0.075 [0.027, 0.123] (<0.01)	0.016 (<0.01)
Ct.Th + Ct.vBMD + EndoTb.vBMD	0.925 [0.881, 0.968]	0.081 [0.027, 0.134] (<0.01)	0.093 (<0.05)
vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.919 [0.874, 0.964]	0.075 [0.024, 0.125] (<0.01)	0.086 (<0.05)
Shape + vBMD	0.924 [0.880, 0.968]	0.080 [0.034, 0.126] (<0.001)	0.093 (<0.01)
Shape + Ct.Th	0.907 [0.858, 0.955]	0.062 [0.002, 0.123] (<0.05)	0.083 (0.076)
Shape + Ct.vBMD	0.852 [0.789, 0.914]	0.007 [−0.064, 0.079] (NS)	−0.043 (NS)
Shape + EndoTb.vBMD	0.925 [0.882, 0.969]	0.081 [0.027, 0.136] (<0.01)	0.069 (0.091)
Shape + vBMD + Ct.Th	0.923 [0.879, 0.967]	0.079 [0.030, 0.127] (<0.01)	0.108 (<0.01)
Shape + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.915 [0.868, 0.961]	0.071 [0.022, 0.119] (<0.01)	0.083 (<0.05)
Shape + vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.927 [0.884, 0.970]	0.083 [0.038, 0.129] (<0.001)	0.090 (<0.05)

50 controls - 93 hip fracture cases

NS = Non-significant

Features in *italic* font are SMPM features, i.e. *m* principal component scores, where *m*=number of training subjects (minus 1) in the *k* fold

Table 3.

Hip Fracture Discrimination and Comparisons to Neck.aBMD with the Inclusion of Demographic and Clinical Variables

Initial Features in the LASSO Model + age + height + weight	AUC [95% CI]	AUC Difference [95% CI] (p-Value)	IDI Index (p-Value)
Neck.aBMD	0.872 [0.814, 0.930]		
Total.aBMD	0.869 [0.810, 0.927]	−0.003 [−0.025, 0.019] (NS)	−0.022 (NS)
Neck.vBMD	0.863 [0.804 0.923]	−0.008 [−0.027, 0.010] (NS)	−0.032 (<0.05)
Total.vBMD	0.863 [0.803, 0.923]	−0.009 [−0.029, 0.011] (NS)	−0.038 (<0.05)
Shape	0.820 [0.751, 0.889]	−0.052 [−0.128, 0.024] (NS)	−0.246 (<0.001)
vBMD	0.917 [0.871, 0.963]	0.045 [−0.001, 0.090] (0.053)	0.039 (NS)
Ct.Th	0.920 [0.876, 0.965]	0.049 [−0.004, 0.101] (0.070)	0.070 (NS)
Ct.vBMD	0.899 [0.848, 0.950]	0.027 [−0.027, 0.081] (NS)	0.028 (NS)
EndoTb.vBMD	0.911 [0.863, 0.959]	0.039 [−0.011, 0.089] (NS)	0.043 (NS)
vBMD + CtTh	0.920 [0.875, 0.965]	0.048 [0.006, 0.090] (<0.05)	0.059 (NS)
Ct.Th + Ct.vBMD + EndoTb.vBMD	0.932 [0.891, 0.973]	0.060 [0.012, 0.108] (<0.05)	0.077 (0.081)
vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.920 [0.876, 0.965]	0.049 [0.002, 0.095] (<0.05)	0.058 (NS)
Shape + vBMD	0.920 [0.876, 0.965]	0.049 [0.002, 0.095] (<0.05)	0.069 (0.078)
Shape + Ct.Th	0.926 [0.882, 0.969]	0.054 [0.003, 0.104] (<0.05)	0.074 (0.084)
Shape + Ct.vBMD	0.864 [0.805, 0.924]	−0.008 [−0.063, 0.048] (NS)	−0.073 (NS)
Shape + EndoTb.vBMD	0.919 [0.874, 0.964]	0.047 [−0.006, 0.101] (0.080)	0.056 (NS)
Shape + vBMD + Ct.Th	0.925 [0.882, 0.969]	0.054 [0.015, 0.092] (<0.01)	0.082 (<0.05)
Shape + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.920 [0.875, 0.965]	0.048 [0.009, 0.088] (<0.05)	0.053 (NS)
Shape + vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.935 [0.895, 0.975]	0.063 [0.026, 0.100] (<0.001)	0.058 (NS)

50 controls - 93 hip fracture cases

NS = Non-significant

Features in *italic* font are SMPM features, i.e. m principal component scores, where m=number of training subjects (minus 1) in the k fold

Age, height, and weight were forced to be part of all LASSO models

Table 4.

Improvement on Hip Fracture Discrimination with Respect to Neck.aBMD (0.844 [0.780, 0.908])

Initial Features in the LASSO Model + Neck.aBMD	AUC [95% CI]	AUC Difference [95% CI] (p-Value)	IDI Index (p-Value)
Shape	0.886 [0.831, 0.940]	0.041 [0.020, 0.071] (<0.01)	0.064 (<0.001)
vBMD	0.907 [0.858, 0.955]	0.062 [0.023, 0.102] (<0.01)	0.178 (<0.001)
Ct.Th	0.908 [0.859, 0.956]	0.064 [0.012, 0.155] (<0.05)	0.230 (<0.001)
Ct.vBMD	0.924 [0.880, 0.968]	0.080 [0.032, 0.128] (<0.01)	0.230 (<0.001)
EndoTb.vBMD	0.914 [0.867, 0.961]	0.070 [0.025, 0.115] (<0.01)	0.218 (<0.001)
vBMD + CtTh	0.911 [0.864, 0.959]	0.067 [0.022, 0.112] (<0.05)	0.204 (<0.001)
Ct.Th + Ct.vBMD + EndoTb.vBMD	0.920 [0.876, 0.965]	0.076 [0.030, 0.123] (<0.01)	0.236 (<0.001)
vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.914 [0.867, 0.961]	0.070 [0.023, 0.117] (<0.01)	0.220 (<0.001)
Shape + vBMD	0.909 [0.861, 0.957]	0.065 [0.016, 0.114] (<0.01)	0.228 (<0.001)
Shape + Ct.Th	0.920 [0.874, 0.965]	0.075 [0.029, 0.122] (<0.01)	0.243 (<0.001)
Shape + Ct.vBMD	0.906 [0.857, 0.955]	0.062 [0.019, 0.105] (<0.01)	0.174 (<0.001)
Shape + EndoTb.vBMD	0.925 [0.881, 0.968]	0.081 [0.035, 0.126] (<0.001)	0.235 (<0.001)
Shape + vBMD + Ct.Th	0.911 [0.863, 0.959]	0.067 [0.022, 0.111] (<0.01)	0.231 (<0.001)
Shape + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.917 [0.871, 0.963]	0.073 [0.030, 0.116] (<0.001)	0.214 (<0.001)
Shape + vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.918 [0.872, 0.963]	0.074 [0.031, 0.116] (<0.001)	0.222 (<0.001)

50 controls - 93 hip fracture cases

NS = Non-significant

Features in *italic* font are SMPM features, i.e. m principal component scores, where m=number of training subjects (minus 1) in the k fold

Neck.aBMD was forced to be part of all LASSO models.

Table 5.

Improvement on Hip Fracture Discrimination with Respect to Neck.aBMD with the Inclusion of Demographic and Clinical Variables (0.872 [0.814, 0.930])

Initial Features in the LASSO Model + Neck.aBMD + age + height + weight	AUC [95% CI]	AUC Difference [95% CI] (p-Value)	IDI Index (p-Value)
Shape	0.887 [0.833, 0.941]	0.015 [−0.020, 0.051] (NS)	−0.002 (NS)
vBMD	0.907 [0.858, 0.956]	0.035 [−0.006, 0.076] (0.093)	0.109 (<0.01)
Ct.Th	0.912 [0.864, 0.959]	0.040 [−0.004, 0.083] (0.073)	0.152 (<0.001)
Ct.vBMD	0.925 [0.881, 0.968]	0.053 [0.014, 0.093] (<0.01)	0.158 (<0.001)
EndoTb.vBMD	0.912 [0.865, 0.960]	0.041 [0.003, 0.078] (<0.05)	0.143 (<0.001)
vBMD + CtTh	0.911 [0.863, 0.958]	0.039 [−0.003, 0.081] (0.069)	0.128 (<0.001)
Ct.Th + Ct.vBMD + EndoTb.vBMD	0.923 [0.879, 0.967]	0.051 [0.010, 0.092] (<0.05)	0.163 (<0.001)
vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.913 [0.866, 0.960]	0.041 [0.002, 0.081] (<0.05)	0.145 (<0.001)
Shape + vBMD	0.909 [0.861, 0.957]	0.037 [−0.007, 0.081] (NS)	0.154 (<0.001)
Shape + Ct.Th	0.925 [0.881, 0.968]	0.053 [0.013, 0.093] (<0.01)	0.171 (<0.001)
Shape + Ct.vBMD	0.906 [0.857, 0.955]	0.034 [−0.001, 0.069] (0.059)	0.106 (<0.001)
Shape + EndoTb.vBMD	0.926 [0.883, 0.969]	0.054 [0.010, 0.098] (<0.05)	0.162 (<0.001)
Shape + vBMD + Ct.Th	0.915 [0.868, 0.961]	0.043 [0.006, 0.080] (<0.05)	0.159 (<0.001)
Shape + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.915 [0.868, 0.971]	0.043 [0.004, 0.081] (<0.05)	0.138 (<0.001)
Shape + vBMD + Ct.Th + Ct.vBMD + EndoTb.vBMD	0.918 [0.872, 0.963]	0.046 [0.010, 0.082] (<0.05)	0.148 (<0.001)

50 controls - 93 hip fracture cases

NS = Non-significant

Features in *italic* font are SMPM features, i.e. m principal component scores, where m=number of training subjects (minus 1) in the k fold

Neck.aBMD, age, height and weight were forced to be part of all LASSO models.