

# A Statistical Model of Shape and Bone Mineral Density Distribution of the Proximal Femur for Fracture Risk Assessment

Tristan Whitmarsh<sup>1</sup>, Karl D. Fritscher<sup>2</sup>, Ludovic Humbert<sup>1</sup>,  
Luís Miguel Del Rio Barquero<sup>3</sup>, Tobias Roth<sup>4</sup>, Christian Kammerlander<sup>4</sup>,  
Michael Blauth<sup>4</sup>, Rainer Schubert<sup>2</sup>, and Alejandro F. Frangi<sup>1</sup>

<sup>1</sup> Center for Computational Imaging & Simulation Technologies in Biomedicine (CISTIB), Universitat Pompeu Fabra (UPF) and CIBER-BBN, Barcelona, Spain

<sup>2</sup> Institute for Biomedical Image Analysis (IBIA), University for Health Sciences, Medical Informatics and Technology (UMIT), Hall in Tirol, Austria

<sup>3</sup> CETIR Centre Mèdic, Barcelona, Spain

<sup>4</sup> Department for Trauma Surgery and Sports Medicine, Innsbruck Medical University, Innsbruck, Austria

**Abstract.** This work presents a statistical model of both the shape and Bone Mineral Density (BMD) distribution of the proximal femur for fracture risk assessment. The shape and density model was built from a dataset of Quantitative Computed Tomography scans of fracture patients and a control group. Principal Component Analysis and Horn's parallel analysis were used to reduce the dimensionality of the shape and density model to the main modes of variation. The input data was then used to analyze the model parameters for the optimal separation between the fracture and control group. Feature selection using the Fisher criterion determined the parameters with the best class separation, which were used in Fisher Linear Discriminant Analysis to find the direction in the parameter space that best separates the fracture and control group. This resulted in a Fisher criterion value of 6.70, while analyzing the Dual-energy X-ray Absorptiometry derived femur neck areal BMD of the same subjects resulted in a Fisher criterion value of 0.98. This indicates that a fracture risk estimation approach based on the presented model might improve upon the current standard clinical practice.

## 1 Introduction

With the rapid advancement of medical imaging technologies, the development of image analysis methods has progressed in accordance. Specifically statistical models have received a great deal of interest and have been used in a wide range of fields for the diagnosis of diseases and detection of symptoms.

Research on statistical models for orthopaedics have mainly focused on modeling the bone shape for reconstructions, which are to be used in therapy planning, and not so much for analysis and diagnosis. In Schuler *et al.* [1], a method is presented to estimate the fracture load using statistical appearance models. This

gives an estimation of the femur strength, which could potentially be used to derive a fracture risk. In current clinical practice, Dual-energy X-ray Absorptiometry (DXA) derived femur neck areal Bone Mineral Density (BMD) is used. This measure however is limited by its two-dimensionality while it has been shown that the 3D distribution of the BMD greatly influences the fracture risk [2]. In Li *et al.* [3], a statistical model was constructed of the proximal femur by image registration and the principal components of this model were evaluated for their discriminatory power. This work however only uses statistical information about the BMD distribution, while it has been shown that also the shape has an influence on fracture risk [4].

In this work, we propose a statistical model of the BMD distribution as well as the shape, whereby, by analyzing the input data consisting of Quantitative Computed Tomography (QCT) scans of fracture patients and a control group, the variation that influences the fracture risk most is determined. This in turn can be used to derive a fracture risk for an unseen patient. The underlying shape and density model is constructed using an intensity based registration process on the QCT scans. Principal Component Analysis (PCA) results in a dimensionality reduction whereby Horn's parallel analysis [5] determines the number of modes of variation to be used for further analysis. As a result, each subject in the model can be said to have a specific set of shape and density model parameters. Using the Fisher criterion, the parameters that best separate the fracture and non-fracture group were determined. Over these parameter values, Fisher Linear Discriminant Analysis (FLDA) was applied, which determined the direction in the feature space that best separates the two classes. The resulting variation is analyzed and compared with previous findings on the shape and BMD distribution with respect to the fracture risk. Finally, the class separation resulting from FLDA was compared to the class separation using the femur neck areal BMD of the same subjects, which is the measurement used in clinical routine to assess the fracture risk.

## 2 Method

### 2.1 Registrations

The model construction depends on accurate intensity based registrations of the QCT volumes. In order to prevent the pelvis area from resulting in misalignments in the registration process, the pelvis area is semi automatically removed from the volumes. In addition, a thresholding is applied, which removes the soft tissue structures that can negatively affect the registration process.

Figure 1 depicts the model construction pipeline. First, a reference volume is chosen based on its regular shape and BMD quality. In this volume, the bone is manually segmented and a regular mesh is constructed from this segmentation. All volumes are subsequently registered to this reference volume by an intensity based similarity registration. This is followed by a multi-scale intensity based Thin Plate Spline (TPS) registration using the mesh vertices as the control points on the reference image. In order to reduce the computation load of

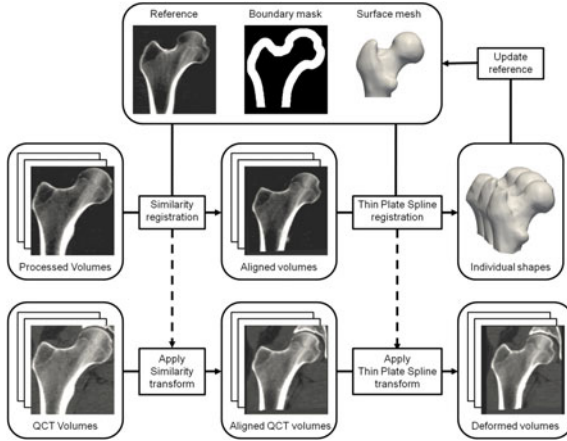


Fig. 1. Model construction pipeline

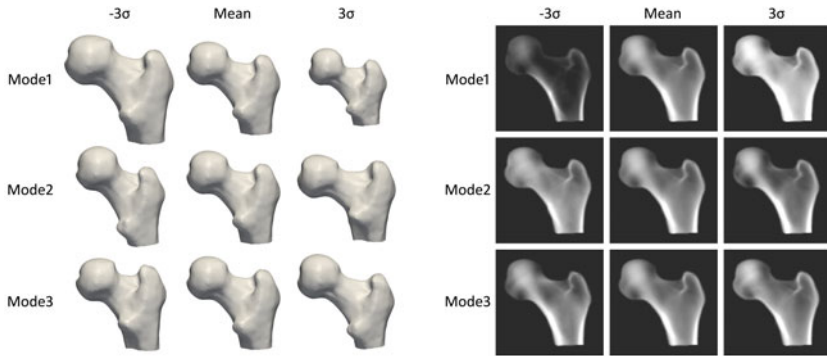
the TPS registration, the registration is restricted to the region of interest by specifying a mask of the bone boundary in the reference volume (Figure 1). The TPS registrations result in the corresponding landmark locations on the target volumes and thus provides the surface mesh for all similarity aligned volumes. All similarity aligned meshes are subsequently scaled to their original size using the uniform scale value resulting from the similarity transform.

To remove some of the bias from the reference selection, the reference volume (and mask) are deformed to the average of the similarity aligned meshes using a TPS transformation defined by the vertices. In a second iteration, the volumes are then registered onto this updated reference while using the average shape for the TPS registrations.

## 2.2 Shape and Density Model

The registrations results in the set of aligned patient specific surface meshes with a vertex correspondence between them. PCA is then applied to the vertices and a new shape  $\mathbf{s}$  can be expressed as the average shape  $\bar{\mathbf{s}}$  and a linear combination of the first  $m$  eigenvectors corresponding to the main modes of variation:  $\mathbf{s} = \bar{\mathbf{s}} + \sum_{i=1}^m \mathbf{p}_i a_i$ . Here  $\mathbf{p}_i$  is the  $i$ -th eigenvector resulting from the singular value decomposition of the covariance matrix and  $a_i$  the corresponding scalar coefficient referred to here as the shape model parameter.

To model the BMD distribution, a final iteration of the registration process is performed to deform all volumes to the same average bone shape, whereafter the resulting similarity and TPS transformation is applied to the unprocessed QCT volume for each subject. This results in shape normalized volumes with a voxel correspondence between them. Since the TPS transformation is defined by landmarks on the bone surface only, the TPS interpolation preserves the internal BMD distribution. PCA is then applied to the voxel densities inside the bone so that a new volume  $\mathbf{v}$  can be expressed as the average volume  $\bar{\mathbf{v}}$  and the



**Fig. 2.** The mean and the first three modes of variation of the shape model (left) and projections of the density model (right), varying 3 standard deviations ( $\sigma$ )

first  $n$  eigenvectors:  $\mathbf{v} = \bar{\mathbf{v}} + \sum_{i=1}^n \mathbf{q}_i b_i$ . Here  $\mathbf{q}_i$  is the  $i$ -th eigenvector and  $b_i$  the corresponding density model parameter. To determine the number of shape model parameters  $m$  and density model parameters  $n$  to retain, Horn’s parallel analysis was used.

### 2.3 Fisher Linear Discriminant Analysis

After the construction of the shape and density model, the model parameter vectors of the femora in both fracture and control group are used as input for FLDA in order to determine the vector in the feature space that best separates these two groups.

Not all parameters resulting from the dimensionality reduction will contribute to a good class separation. Therefore, first a feature selection based on the Fisher criterion is performed. The Fisher criterion for feature  $k$  is defined as:  $F(k) = \frac{\mathbf{S}_B^k}{\mathbf{S}_W^k}$ , where  $\mathbf{S}_B^k$  and  $\mathbf{S}_W^k$  are the  $k$ -th diagonal element of  $\mathbf{S}_B$  and  $\mathbf{S}_W$ , respectively. Here  $\mathbf{S}_B$  and  $\mathbf{S}_W$  are the between-class and within-class scatter matrices and are defined as  $\mathbf{S}_B = (\mathbf{m}_1 - \bar{\mathbf{m}})(\mathbf{m}_1 - \bar{\mathbf{m}})^T + (\mathbf{m}_2 - \bar{\mathbf{m}})(\mathbf{m}_2 - \bar{\mathbf{m}})^T$  and  $\mathbf{S}_W = \sum_{\mathbf{x} \in C_1} (\mathbf{x} - \mathbf{m}_1)(\mathbf{x} - \mathbf{m}_1)^T + \sum_{\mathbf{x} \in C_2} (\mathbf{x} - \mathbf{m}_2)(\mathbf{x} - \mathbf{m}_2)^T$ .  $C_1$  and  $C_2$  are the feature sets of the fracture and non-fracture group,  $\mathbf{m}_1$  and  $\mathbf{m}_2$  are their respective means and  $\bar{\mathbf{m}}$  the overall mean. Each feature is analyzed for their class separation by calculating this Fisher criterion. The features are subsequently sorted and the features beyond the converging point (nearing zero) are discarded.

FLDA is applied to the remaining features to determine the vector that best separates the fracture and control group. FLDA finds the vector  $\mathbf{w}$  that maximizes the total scatter while minimizing the within scatter of the two classes. This is achieved by maximizing the Rayleigh quotient:  $J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}$ .

The resulting direction in the feature space is visualized by generating an instance of the model at the two extremes of this vector corresponding to 3 and -3 standard deviation from mean. Three standard deviations are chosen to

constrain them to statistically valid instances. For both this high and low fracture risk extreme, the density model instance is deformed towards the shape model instance, again using a TPS transformation defined by the shape vertices. These visualizations allow for the assessment of the combined variations of shape and BMD distribution that influences the fracture risk, and these are subsequently compared to current knowledge of hip fracture predictors.

## 2.4 Data

A dataset of CT scans of the pelvis area of 58 fracture patients was collected at the Universitätsklinik für Radiodiagnostik in Innsbruck using the GE Light-Speed VCT Multi Slice CT device (GE Healthcare, Chalfont St. Giles, UK). The patients were all female with an average age of  $79 \pm 10$  years and all had suffered a proximal femur fracture. From this set the contralateral un-fractured femurs were taken, which can be justified by research on proximal femur symmetry [6], and this set defined the high fracture risk group. In addition, the European forearm phantom [7] was included into every scan acquisition to convert the CT scans to QCT volumes.

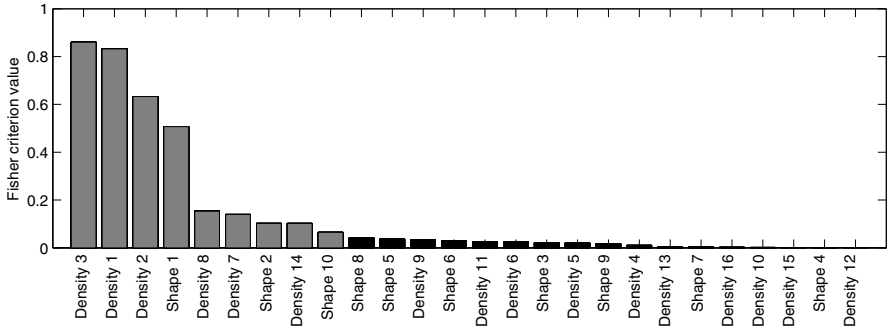
Besides the fracture group, a group of 58 (all female) patients was collected at the CETIR Medical Center (Barcelona, Spain) using the Philips Gemini GXL 16 system (Philips Healthcare, Best, The Netherlands). These patients had a lower average age of  $55 \pm 12$  years, to represent a control group with a low fracture risk and a normal BMD distribution. The CT scans were calibrated using the Mindways calibration phantom (Mindways Software Inc., Austin, TX, United States).

Different studies have shown that the scan device has little influence on the density calibration [8]. However, the European forearm phantom calibrates the volumes to a hydroxyapatite (HA) density, whereas the Mindways phantom is constructed of  $K_2HPO_4$ . Previous research shows that one calibration material is highly correlated to another and can be converted using a linear transformation [9]. To get the conversion formula for the phantoms in this work, both phantoms were scanned together using the Siemens scanner. This resulted in the following conversion formula:  $y = 1.1053x - 17.788$  ( $R^2 = 0.9998$ ). This way, the scans with the European forearm phantom were calibrated to relate to  $K_2HPO_4$  densities.

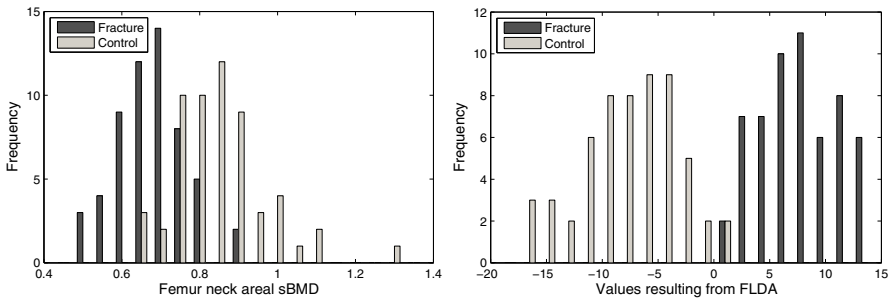
Femur neck BMD measurements were performed on the DXA scans of all subjects. For the fracture group patients, the Hologic Discovery W bone densitometry system (Hologic Inc., Bedford, MA, United States) was used while for the control group the scans were performed using the GE Healthcare's Lunar iDXA scanner (GE Healthcare, Chalfont St. Giles, UK). The femur neck areal BMD values were subsequently converted to standardized BMD (sBMD) [10].

## 3 Results

Figure 2 shows the first three modes of variation of the resulting shape and density model. Horn's parallel analysis resulted in 10 shape model parameters



**Fig. 3.** A bar graph of the features with the corresponding Fisher criterion value, sorted in descending order. The gray bars indicate the features selected for further analysis while the features with the black bars were discarded.

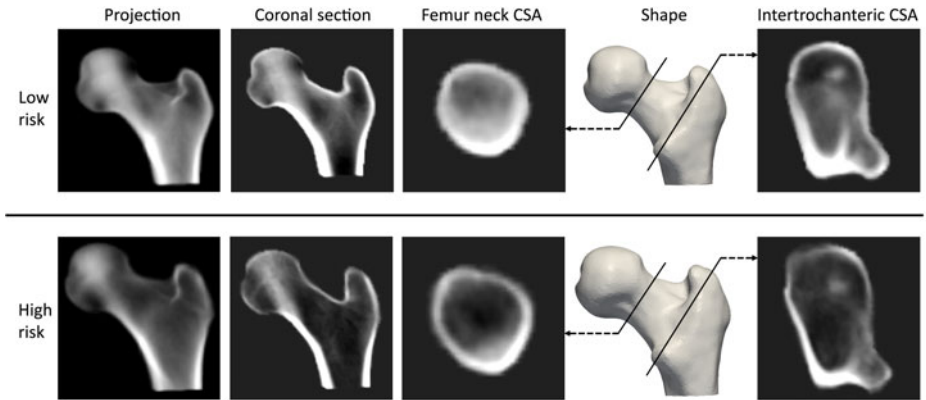


**Fig. 4.** The histograms corresponding to the femur neck areal sBMD (left) and the values resulting from FLDA (right) for both fracture and control group.

and 16 density model parameters to form the feature space. The feature selection method using the Fisher criterion resulted in 3 shape model parameters and 6 density model parameters to be used, as shown in Figure 3. Applying FLDA to the values of these parameters of the subjects in the model resulted in a class separation with a Fisher criterion value of 6.70. In comparison, the femur neck areal sBMD values separates the fracture and control group with a Fisher criterion value of 0.98.

For each subject in the model, the corresponding shape and density model parameters can be projected onto the vector resulting from FLDA. This represents a single value measure for the fracture risk. In Figure 4, the histogram corresponding to these values are shown for both fracture and control group. The same histogram is given for the femur neck areal sBMD values as a comparison.

In Figure 5, the high and low fracture risk extremes with respect to the mean are visualized, which show the difference in shape and BMD distribution between high and low fracture risk femurs as described by the model.



**Fig. 5.** Model instances at the low fracture risk (top) and high fracture risk (bottom) extreme. From left to right: the projection, the coronal cross section, the femur neck cross-sectional area (CSA), the shape model instance and the intertrochanteric CSA.

## 4 Discussion and Conclusions

This work presents a shape and density model of the proximal femur for fracture risk estimation. A dimensionality reduction and feature selection method is used to extract the shape and density model parameters that best separate the fracture and control group. FLDA is subsequently used to find the direction in this parameter space of greatest separation.

In Figure 5 it can be seen that the volumetric BMD and the cortical thickness is one of the main discriminators between fracture and non-fracture femora according to the presented model, which is in accordance with recent findings [2]. The visualization of the femur neck and intertrochanteric cross-sectional area (CSA) also reflects previous observations where a significantly smaller cortical CSA, and a larger trabecular CSA is associated with a higher fracture risk [4]. Regarding the shape, a larger hip axis length is associated with a higher fracture incidence [4], which is also reflected in the presented model. The neck-shaft angle has been shown to be significantly larger for fracture patients [4], and this is captured by the model through the second shape model parameter (Figure 2 and 3). These similarities indicate that the presented model captures the differences in shape and BMD distribution between high and low fracture risk subjects and can thus be used in a computer aided diagnosis system.

The Fisher criterion value resulting from FLDA shows that the proposed model better separates the fracture and control group than DXA derived femur neck areal sBMD with a value of 6.70 as opposed to 0.98. The improved class separation can also be seen in the histograms of Figure 4. This indicates that a model-based fracture risk estimation approach might improve upon the current standard clinical practice.

**Acknowledgments.** The research leading to these results has received funding from the ERDF Operational Programme of Catalonia 2007-2013, through the 3D-FemOs project (exp.VALTEC09-02-0012) and the EU FP7 programme under grant agreement nr 269909 ( MySPINE project). A.F. Frangi is partially funded by the ICREA-Academia programme and T. Whitmarsh is supported by the Instituto de Salud Carlos III through PFIS predoctoral grant (FI09/00757).

## References

- Schuler, B., Fritscher, K.D., Kuhn, V., Eckstein, F., Link, T.M., Schubert, R.: Assessment of the individual fracture risk of the proximal femur by using statistical appearance models. *Medical Physics* 37(6), 2560–2571 (2010)
- Bousson, V., Adams, J., Engelke, K., Aout, M., Cohen-Solal, M., Bergot, C., Hagnauer, D., Goldberg, D., Champion, K., Aksouh, R., Vicaut, E., Laredo, J.: 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 (2010)
- Li, W., Kornak, J., Harris, T., Lu, Y., Cheng, X., Lang, T.: Hip fracture risk estimation based on principal component analysis of QCT atlas: a preliminary study. In: *Medical Imaging 2009: Biomedical Applications in Molecular, Structural, and Functional Imaging*, p. 72621M. SPIE, San Jose (2009)
- Ito, M., Wakao, N., Hida, T., Matsui, Y., Abe, Y., Aoyagi, K., Uetani, M., Harada, A.: Analysis of hip geometry by clinical CT for the assessment of hip fracture risk in elderly Japanese women. *Bone* 46(2), 453–457 (2010)
- Horn, J.: A rationale and test for the number of factors in factor analysis. *Psychometrika* 30(2), 179–185 (1965)
- Faulkner, K., Genant, H., McClung, M.: Bilateral comparison of femoral bone density and hip axis length from single and fan beam DXA scans. *Calcified Tissue International* 56, 26–31 (1995)
- Ruegsegger, P., Kalender, W.A.: A phantom for standardization and quality control in peripheral bone measurements by PQCT and DXA. *Physics in Medicine and Biology* 38(12), 1963–1970 (1993)
- Fujii, Y., Tsunenari, T., Tsutsumi, M., Miyauchi, A., Yamada, H., Fukase, M., Yoshimoto, Y., Okuno, Y., Kusakabe, H., Miyoshi, K., Fukunaga, M., Morita, R., Fujita, T.: Quantitative computed tomography: Comparison of two calibration phantoms. *Journal of Bone and Mineral Metabolism* 6, 17–20 (1988)
- Suzuki, S., Yamamuro, T., Okumura, H., Yamamoto, I.: Quantitative computed tomography: comparative study using different scanners with two calibration phantoms. *British Journal of Radiology* 64(767), 1001–1006 (1991)
- Genant, H.K., Grampp, S., Glüer, C.C., Faulkner, K.G., Jergas, M., Engelke, K., Hagiwara, S., van Kuijk, C.: Universal standardization for dual X-ray absorptiometry: Patient and phantom cross-calibration results. *Journal of Bone and Mineral Research* 9(10), 1503–1514 (1994)