Biomech Model Mechanobiol (2011) 10:529–538
DOI 10.1007/s10237-010-0253-3

ORIGINAL PAPER

Assessment of the 3-D shape and mechanics of the proximal femur using a shape template and a bone mineral density image

Sami P. Väänänen · Hanna Isaksson · Petro Julkunen · Joonas Sirola · Heikki Kröger · Jukka S. Jurvelin

Received: 9 June 2010 / Accepted: 17 August 2010 / Published online: 1 September 2010

Abstract Measurement of bone mineral density (BMD) by DXA (dual-energy X-ray absorptiometry) is generally considered to be the clinical golden standard technique to diagnose osteoporosis. However, BMD alone is only a moderate predictor of fracture risk. Finite element analyses of bone mechanics can contribute to a more accurate prediction of fracture risk. In this study, we applied a method to estimate the 3D geometrical shape of bone based on a 2D BMD image and a femur shape template. Proximal femurs of eighteen human cadavers were imaged with computed tomography (CT) and divided into two groups. Image data from the first group ($N = 9$) were applied to create a shape template by using the general Procrustes analysis and thin plate splines. This template was then applied to estimate the shape of the femurs in the second group ($N = 9$), using the 2D BMD image projected from the CT image, and the geometrical errors of the shape estimation method were evaluated. Finally, finite element analysis with stance loading condition was conducted based on the original CT and the estimated geometrical shape to evaluate the effect of the geometrical errors on the outcome of the simulations. The volumetric errors induced by the shape estimation method itself were low (<0.6%). Increasing the number of bone specimens used for the template decreased the geometrical errors. When nine bones were used for the template, the mean distance difference (±SD) between the estimated and the CT shape surfaces was 1.2 ± 0.3 mm, indicating that the method was feasible for estimating the shape of the proximal femur. Small errors in geometry led systematically to larger errors in the mechanical simulations. The method could provide more information of the mechanical characteristics of bone based on 2D BMD radiography and could ultimately lead to more sensitive diagnosis of osteoporosis.

Keywords Proximal femur · Shape estimation · Finite element · Bone strength · Computed tomography

S. P. Väänänen (✉) · H. Isaksson · J. S. Jurvelin
Department of Physics and Mathematics,
University of Eastern Finland, PO Box 1627,
70211 Kuopio, Finland
e-mail: sami.vaananen@uef.fi

P. Julkunen
Department of Clinical Neurophysiology,
Kuopio University Hospital, PO Box 1777, 70211 Kuopio, Finland

P. Julkunen
Department of Clinical Radiology, Kuopio University Hospital,
PO Box 1777, 70211 Kuopio, Finland

J. Sirola · H. Kröger
Bone and Cartilage Research Unit, University of Eastern Finland,
PO Box 1627, 70211 Kuopio, Finland

1 Introduction

Current osteoporosis diagnosis is primarily based on the measurement of bone mineral density (BMD), using dual-energy X-ray absorptiometry (DXA) of the hip or the lumbar spine. Areal BMD is a good measure of bone density but only a moderate predictor of fracture risk. It has been shown that only about 30% of all low energy fractures can be explained by changes in BMD alone (Pasco et al. 2006). Bone quality can be summarized as all characteristics that affect the bone's resistance to fracture, e.g., bone mineral density, architecture, mechanical strength, geometry, turnover and mineralization. Therefore, novel methods that can describe bone properties better than BMD are needed.

In vivo measurements of bone strength are not feasible. However, previous studies suggest that finite element (FE) analysis of the bone mechanics can account for at least 20%

more of the variance in strength of the proximal femur than DXA alone (Cody et al. 1999). In these studies, FE modelling was based on 3D quantitative computed tomography (QCT) images of the femur (Keyak et al. 1998, 2001; Cody et al. 1999; Keyak and Falkinstein 2003; Wirtz et al. 2003; Edwards et al. 2008). However, QCT imaging is more expensive and results in more significant radiation doses to the patient compared with DXA. Therefore, a method that can assess bone strength accurately without the need of 3D CT images is desired.

Three-dimensional computational modelling of the proximal femur using DXA is not straightforward because DXA images lack the information in the anterior-posterior direction. To overcome this restriction, bone templates with information on general 3D shape and architecture of the femur may be adapted to create individual, subject-specific 3D shape models. Previously, the approaches to assess the 3D shape have been based on several radiographic images (Laporte et al. 2003; Kolta et al. 2005; Filippi et al. 2007). Only one recent technique has approximated the 3D shape of the proximal femur using a single X-ray image and a shape template (Langton et al. 2009b). This method could re-create the outer geometry of the femur with a reasonable accuracy. However, more detailed analyses are needed to evaluate the accuracy of the shape estimation methods, e.g. to determine where the greatest estimation errors occur concerning the proximal femur. It is also important to know how these local geometrical errors translate into errors in the estimated mechanical behaviour of the proximal femur. Using the estimated shape, Langton et al. (2009a) conducted FE analysis of bone strength with encouraging results. However, it remains unknown how similar the results are for the FE models derived from a single 2D X-ray image using the estimated shape, as compared to models derived directly from CT images.

The objective of this study was to develop a method to estimate the 3D bone shape, based on a 2D BMD image and a shape template of the proximal femur. Further, we characterize the accuracy of the shape estimation by comparing the original 3D CT image with the estimated 3D shape within the whole proximal femur and specifically in different anatomical regions (head, neck, trochanter and shaft). As a final step, we evaluate how the geometric errors translate into errors in mechanical characteristics of femur using FE analysis.

2 Material and methods

2.1 Samples and computed tomography (CT)

Proximal femurs from 18 human cadavers were used in this study. The samples originated from 7 right and 11 left femurs from 16 men and 2 women, age range of 26–82 years (mean 52 ± 15 years). None of the cadavers had any pre-existing conditions that might affect bone metabolism. Ethical permission was granted from the National Authority for Medicolegal Affairs (TEO, 5783/04/044/07). Each femur was imaged with a clinical CT scanner (Siemens Definition AS64, Siemens AG, Erlangen, Germany) using 0.6- mm isotropic voxel size. During scanning, a CT calibration phantom was used to determine the dipotassium phosphate (K_2HPO_4) content of the femur (Mindways Software Inc., Austin, TX, USA). The phantom included five reference materials calibrated against known liquid K_2HPO_4/water solutions. After scanning, linear regressions were calculated between Hounsfield unit (HU) values of reference materials and K_2HPO_4, and HU values of the reference materials and water content. The water-free K_2HPO_4 content was obtained after the regression of water was subtracted from the original K_2HPO_4 regression.

2.2 Shape assessment

First, all CT images of the left femurs were mirrored in the anterior-posterior direction, in order to have a complete sample set from the right side. Secondly, bone tissue was segmented from other tissues using a constant threshold value of 1,285 HU. After thresholding, gaps inside the outer geometrical shape of the femurs were filled to get one uniform femur image. Additionally, the CT images were re-aligned using 3D pattern recognition. The femurs were aligned by finding the centre of the most distal part of each femoral shaft, the centre of the most proximal part of the major trochanter and the centre of each femur head. Thereafter, the centres were used to align the femurs in the frontal plane.

In shape estimation, three 2D map images were extracted from each CT image set, i.e. *offset* and d*epth* images (Fig. 1), and *BMD* image (Fig. 2). In the offset image, each pixel defines the distance from the posterior surface of the proximal femur to the frontal reference plane at an arbitrary distance from the femur. In the depth image, each pixel defines

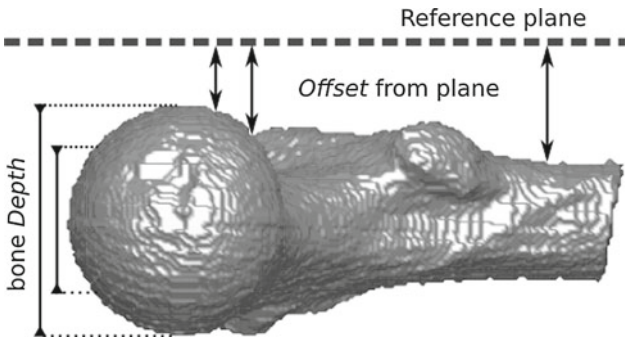

Fig. 1 Definition of offset and depth. The offset is the distance from the posterior surface of the proximal femur to the frontal reference plane at an arbitrary distance from the femur. The depth is the femur thickness in the posterior–anterior direction

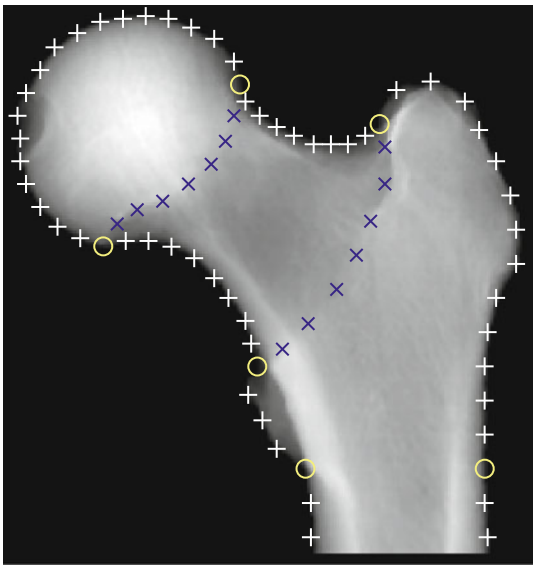

Fig. 2 Bone mineral density (BMD) image of the proximal femur. The BMD image is based on the projection of the CT images in the anterior-posterior direction, and landmarks used in shape estimation are marked. *Yellow circles* are the main anatomical landmarks and *blue crosses* are mathematical landmarks which separate the head and neck areas and the neck and trochanter areas. *White pluses* are pseudo-landmarks located equally spaced on the boundary of the femur image based on the anatomical landmarks

the femur thickness in the posterior–anterior direction. The BMD image is a projection image of the proximal femur where the voxel values are summed in the posterior–anterior direction, essentially resembling a normal X-ray image (Langton et al. 2009b).

A landmark is a point of correspondence on each object that matches between and within a population (Dryden and Mardia 1998). In this work, three types of landmarks (Dryden and Mardia 1998) were defined for each femur, based on the BMD images (Fig. 2).

1. Anatomical landmarks, i.e. landmarks bounded to a specific femoral detail, were positioned manually.
2. Mathematical landmarks, i.e. landmarks that are located according to a mathematical or geometrical property, were positioned semi-automatically to boundaries of different femoral parts.
3. Pseudo-landmarks, i.e. landmarks with locations defined based on the anatomical landmarks, were positioned automatically on the outer boundary of the BMD femur image.

Two methods were applied in the shape estimation. General Procrustes analysis (GPA) is a method which removes affine (i.e. translation, scaling and rotation) transformations between landmark sets (Gower 1975). Thin plate splines (TPS) is a method to deform images based on changed landmark configurations (Bookstein 1989). The inputs for TPS include the image that will be deformed, the landmarks from the original image and the landmarks in the output position, i.e. the locations where they should move to in an image.

CT-scanned and pre-processed femurs were divided randomly into two groups of nine specimens each. The specimens in the first group were used to create the templates, while those in the second group were used to predict the 3D bone shape and to assess the accuracy of the method (Fig. 3, left side). The template creation was started by removing the affine transformations between landmark sets using GPA, after which the mean position of every landmark was calculated. Then, offset and depth images of each template bone were deformed to the shape described by the mean landmark positions using TPS. The mean of the offset and depth images were calculated pixel-by-pixel, after which mean offset and depth images were obtained. These two images constitute the femur template, which describes the mean 3D shape of the femur (Langton et al. 2009b).

After the template was created, it was deformed to the shape of the femurs in the second group to estimate their 3D shape (Fig. 3, right side). The landmarks of the estimated femurs were found from the BMD images of each femur, and TPS was used to deform the offset and depth images of the template to the shape of the second group of bones based on these landmarks. An estimated 3D model of the bone was obtained by combining the offset and depth images. The BMD distribution of the model was obtained from the BMD image. This means that the bone density of the model varies in the medial–lateral and superior-inferior directions of the bone but remains constant in the anterior-posterior direction (Fig. 3, bottom right corner). All the image pre-processing and shape estimation were conducted in Matlab (v. 7.6, Mathworks, Inc., Natick, MA).

The ability of the described method to re-create and estimate the shape of the femurs was evaluated by calculating the volumetric difference, mean distance difference and maximum distance difference between the original and re-created shapes. The volumetric difference was calculated by subtracting the volume of the estimated femur shape from the volume of the original femur shape. The mean distance difference was determined by calculating the average distance from every surface voxel of the original femur to the nearest surface voxel of the estimated shape when the original and estimated shapes were co-registered. The maximum distance difference was determined similar to the mean distance difference, but instead of calculating the average of the distances between the nearest surface points the maximum value was chosen. In the medial–lateral and the superior-inferior directions, the original and estimated shapes were superimposed automatically during the shape estimation. In the anterior-posterior direction, the original and estimated shapes were co-registered by calculating cumulative sums of the voxels and then by moving the shapes so that the locations of the half

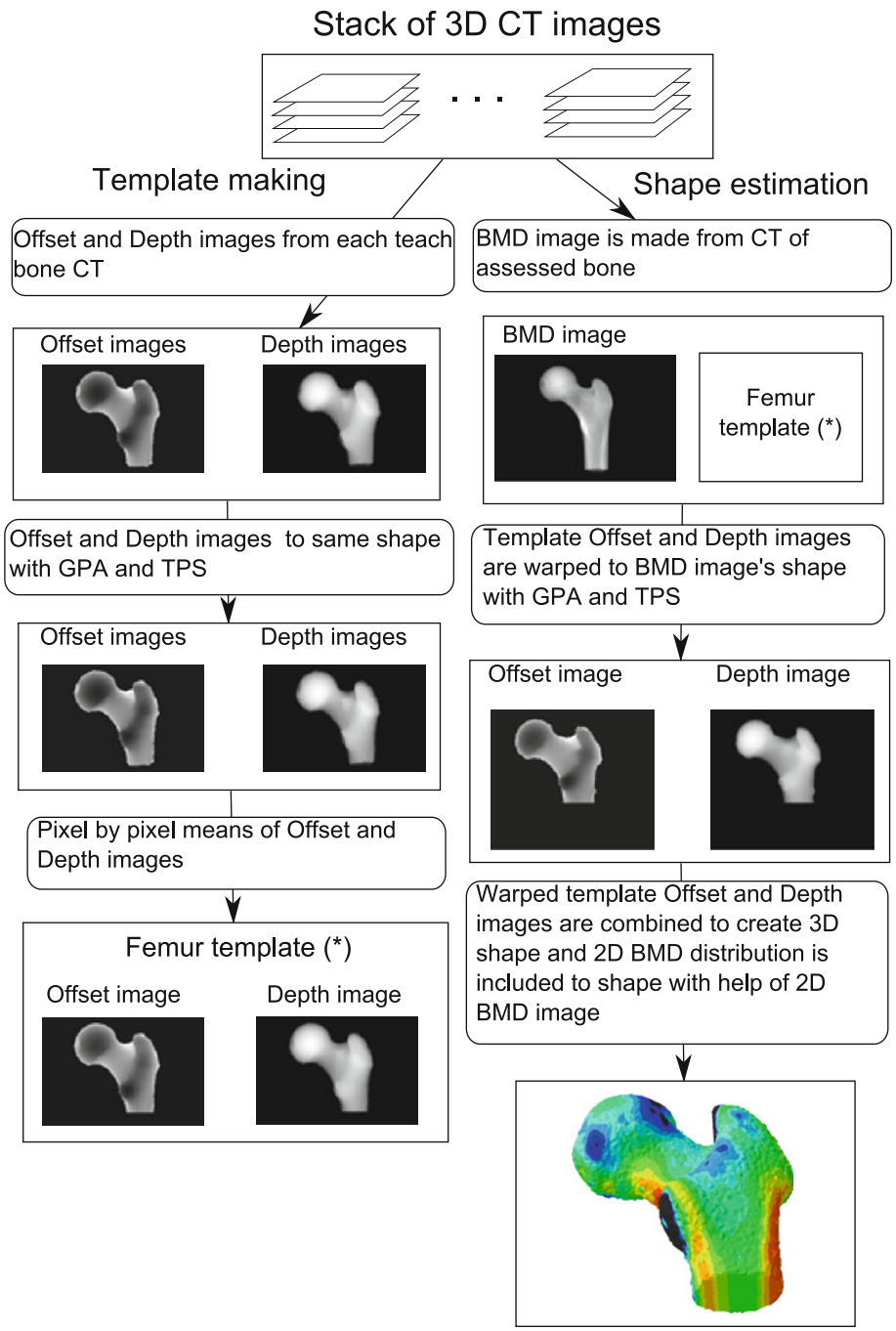

Fig. 3 A flowchart describing the shape estimation procedure. On the *left side*, the creation of the femoral template is described. On the *right side*, the deformation of the femur template to the estimated shape is described. The creation of the template includes using offset and depth images, general Procrustes analysis (GPA) and thin plate splines (TPS) to create the femur template (*). The shape estimation uses the femur template (*) and the BMD image to create the final shape estimate. This shape estimation is used to create the 3D FE model with two-dimensional distribution in BMD values (BMD distribution is constant in anterior-posterior direction)

maximums in the cumulative sums were equalized. These three shape parameters were calculated for the whole proximal femur and for four anatomical parts separately, i.e. the head, neck, trochanter and shaft which were defined on the basis of the landmarks used in template making and shape estimation.

The effect of the number of bones used to create the shape estimation template on the estimation accuracy was investigated by increasing the number of template bones from 1 to 9. First, the number of bones for the template was chosen. Then, the bones for a template were chosen randomly from the set of nine femurs and this was repeated nine times, so that the cases when all the bones were the same as in an earlier case were not counted. This procedure resulted in 73 different templates.

2.3 Finite element modelling

To estimate the resulting differences in the simulated mechanical function, a 3D FE model was created for both the

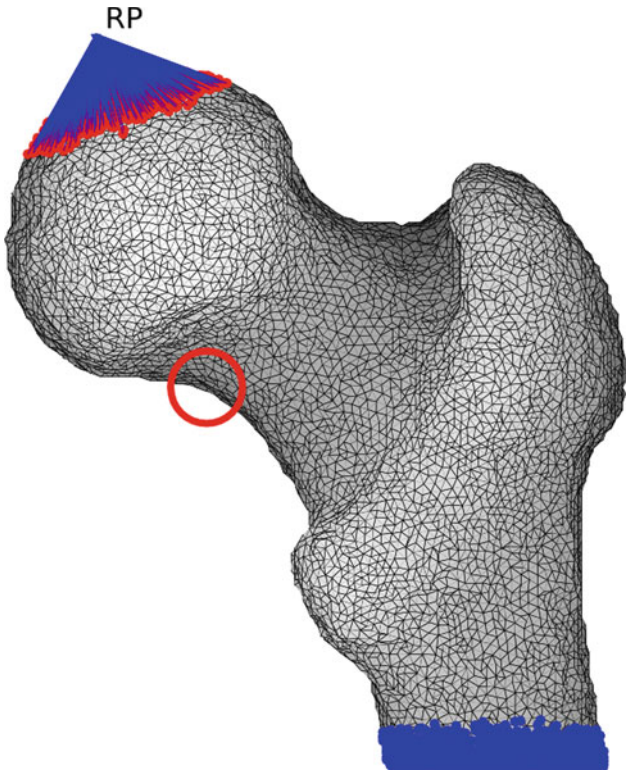

Fig. 4 The FE model used for the strength analysis of the femur. The reference point (RP) to which the force is applied is located above the femoral head. The RP is connected to the nodes on the femoral head (*red dots*) using kinematic coupling. The nodes in the femur shaft, indicated by *blue dots*, are constrained in every direction. The *red circle* on the femur neck indicates the ROI used in the strength analysis

original CT shape and the estimated geometrical shape of each femur. The FE model was generated with tetrahedral mesh in Matlab using an open source meshing software (Fang and Boas 2009). Each model had about 65,000 elements and 15,000 nodes. The FE model was thereafter imported into Abaqus (v 6.8, Simulia, Dassault Systems, Providence, RI, USA).

The nodes in the most distal part of the femoral shaft were constrained in all directions. A node was constrained when it was less than 5 mm from the most distal node in the direction of the shaft axis (Fig. 4). Before mesh creation, each femoral shaft of the models was cut 18 mm below the lowest part of the minor trochanter. The length was chosen based on the shortest cadaver shaft. An external force was applied to the femoral head to simulate stance (Bergmann et al. 2001). The forces were 2.25 and 0.50 times the body-weight in the superior-inferior and medial–lateral loading directions, respectively (Bergmann et al. 2001). The load was applied to a reference point, and then distributed through kinematic coupling to all loading nodes. To mimic the acetabulum, a node was defined as a loading node if the distance from it to the top-most point of the femur head was less than 6 mm in the proximal–distal direction when the femur was rotated 20° around the anterior-posterior axis (Cody et al. 1999) (Fig. 4).

The Hounsfield value of each element was defined from the CT or estimated shape image filtered with 3 × 3 × 3-element averaging mask at coordinates that corresponded to the element centre. The Hounsfield values were mapped to the K_2HPO_4 values using the phantom calibration from the CT imaging, and the K_2HPO_4 values were converted to Young's modulus (Morgan et al. 2003).

The models based on the CT image and the estimated shape were compared with each other according to the following parameters: stiffness, mean of first principal strain, mean of first principal stress and mean of Von Mises stress. The stiffness was defined as the loading force divided by the average displacement of the loading nodes. The mean first principal strain, first principal stress and von Mises stress were calculated in the whole proximal femur, and in the head, neck, trochanter and shaft separately. Furthermore, they were calculated in a smaller region of interest (ROI) in the femoral neck (Fig. 4), where high localized stresses and strains have been found in previous studies (Keyak et al. 1998). The position of the ROI was bound to the landmarks to keep its anatomical position identical between the femurs.

Finally, the differences between the original CT-based model and the model based on the estimated shapes were analysed with two different distributions of Young's modulus. In the first case, the Young's modulus was constant (10 GPa) in both models, to determine the errors caused solely by the shape estimation. In the second case, a 2D distribution of Young's modulus was applied, which means that the values of Young's modulus changed in the medial–lateral and superior-inferior directions but remained constant in the anterior-posterior direction in both models. This analysis determined the errors induced when the most developed shape and Young's modulus distribution estimation method based on one BMD image were used. All simulations were performed with Abaqus using linear elastic material behaviour.

2.4 Statistical analyses

The Kruskal–Wallis test was used to determine whether the estimation accuracy of the model was dependent on the number of bones used to create the template. Bland–Altman plots were applied to determine the agreement of parameter values between the original CT- and estimated shape-based models of the whole proximal femur when Young's modulus was varied in 2D (Bland and Altman 1986). Statistical tests were conducted using SPSS software (v. 14.0, SPSS Inc., Chicago, IL, USA).

Table 1 The methods ability to re-create the shape in different regions of the proximal femur (mean ± SD, based on 9 femurs)

	Whole bone	Head	Neck	Trochanter	Shaft
Volumetric difference (%)	0.32 ± 0.13	0.11 ± 0.08	0.09 ± 0.05	0.62 ± 0.24	0.19 ± 0.18
Mean distance difference (μm)	23 ± 10	7 ± 5	6 ± 3	40 ± 20	10 ± 8
Max distance difference (mm)	2.7 ± 0.9	0.8 ± 0.3	1.3 ± 0.7	2.6 ± 0.9	1.0 ± 0.6

The errors of the re-creation of the shape were small. The largest errors were located on the trochanter area

3 Results

3.1 Shape assessment

The ability to re-create the original shape was first evaluated in order to quantify the errors induced by the method itself, i.e. the errors resulting from taking one femur, creating a shape template and re-creating the same femur. The resulting volumetric difference was less than 0.7% and the mean distance difference was less than 50 μm in all anatomical parts (Table 1). The maximum distance difference ranged from 0.7 to 2.6 mm, with the maximum error found in the trochanter area (Table 1).

Secondly, the effect of the number of bones used to create the shape estimation template was investigated, and the accuracy was evaluated by increasing the number of template bones from 1 to 9. For each template, the shapes of nine proximal femurs were estimated, and the accuracy and difference between the original CT shape and the estimated shape were evaluated qualitatively (Fig. 5). Based on the contours of the CT and the estimated shape, the method achieved better accuracy in the frontal plane than in the anterior-posterior plane (Fig. 5). The mean volumetric difference and the mean and maximum distance differences decreased slightly as the number of template bones was increased (Fig. 6). The volumetric difference was 11 ± 4% for the whole bone, and it did not decrease significantly after the number of bones used to create the template was increased. However, the mean and maximum distance differences decreased significantly ($p < 0.01$) when the number of bones used to create the template was increased. For example, in the neck area, the mean and maximum distance differences decreased ($p < 0.01$) from 1.5 to 1.2 mm and from 6.4 to 4.7 mm, respectively (Fig. 6).

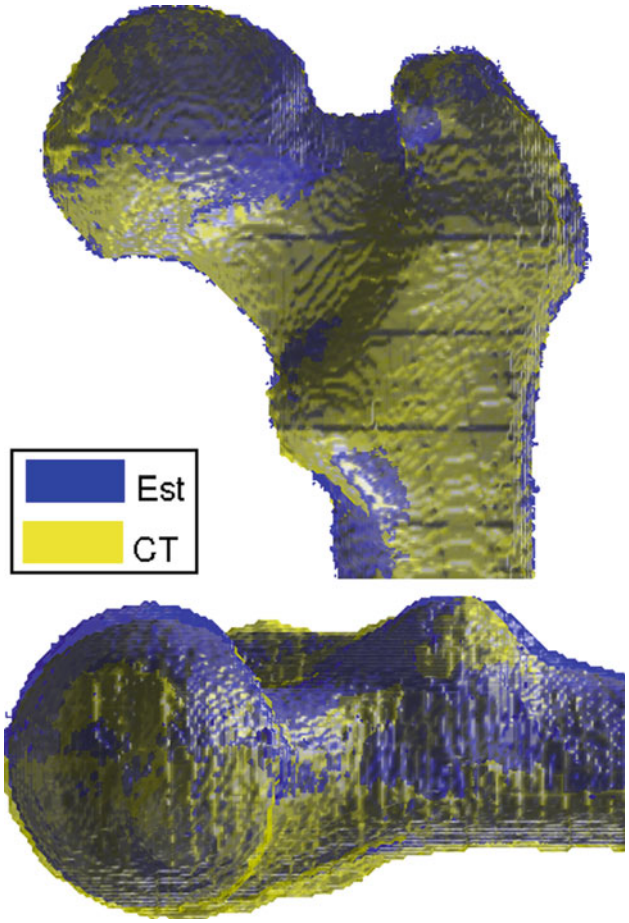

Fig. 5 Original CT image (representative 3D rendering) of the femur (*yellow*) and the estimated shape (*blue*) from posterior–anterior and medial–lateral views. In general, the shape estimation was more accurate in the medial–lateral and superior-inferior directions than in the posterior–anterior direction. The largest errors in shape estimation were located in the area of the major and minor trochanter

3.2 Modelling

The FE modelling was based on the estimated shapes from a shape template of nine bones. The difference in parameter values between the estimated shape and the CT-based models were normalized to the CT-based modelling data, and linear correlations (r^2) between the estimated shape and the CT-based models were evaluated for each result parameter. When using a constant Young's modulus, the difference between the stiffness was 12 ± 4% and when using a non-constant (image-based) Young's modulus the difference was 31 ± 25%. The correlation coefficients were 0.75 and 0.83, respectively ($p < 0.01$).

Normalized differences for the means of the first principal strain, first principal stress and von Mises stress were calculated in the regions of the whole model, head, neck trochanter and the specific ROI in the femoral neck (Fig. 7). When a constant Young's modulus was used, the normalized errors in all modelling parameters were similar in each ana-

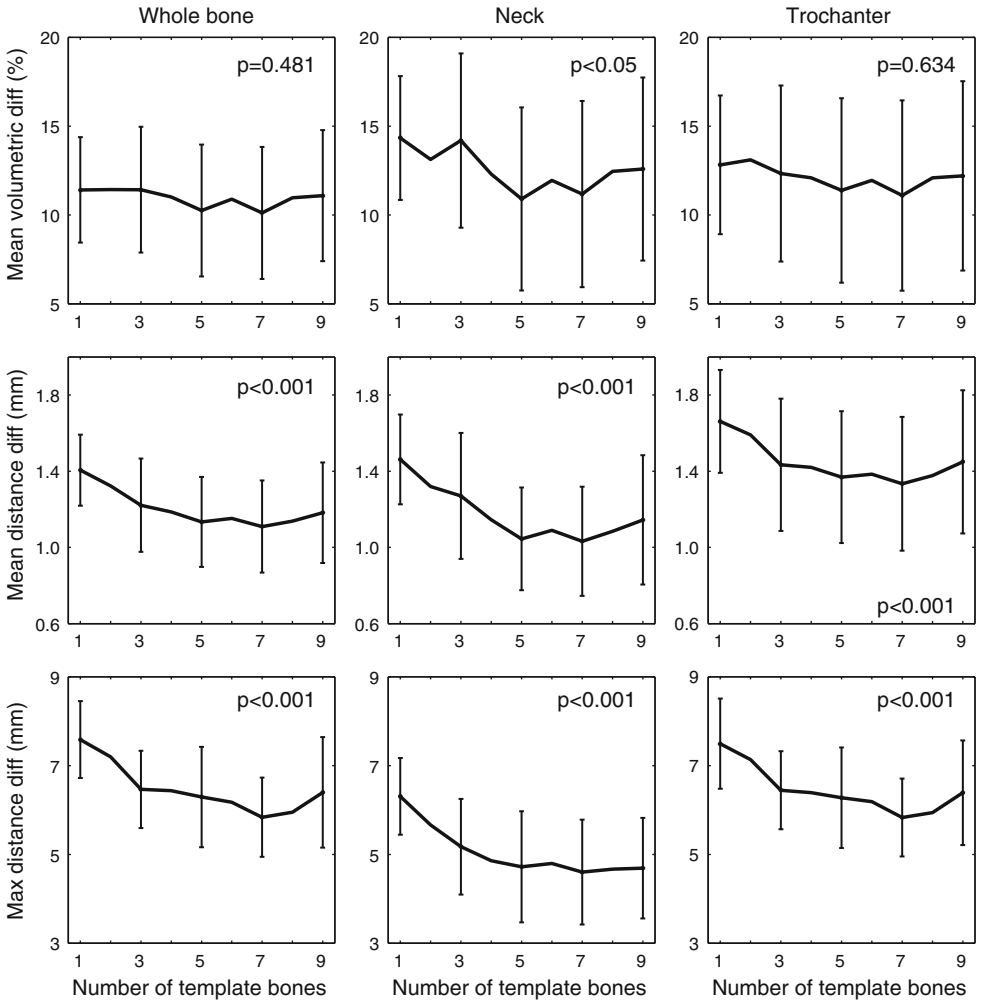

Fig. 6 The effect on accuracy of shape estimation when the number of bones (1–9) used to create the shape template was increased. The *rows* show the change in mean volumetric difference and in mean and maximum distance difference for the whole proximal femur and the anatomical regions of the femoral neck and the trochanter. The mean and standard deviations are presented. For the whole bone, the volumetric difference did not change significantly when the number of template bones was increased, whereas the mean and max distance differences decreased significantly ($p < 0.01$) when the number of template bones was increased. In the neck area, the mean distance difference decreased from 1.45 to 1.15 mm and the maximum distance difference fell from 6.4 to 4.7 mm. Significant differences based on the Kruskal–Wallis test are indicated

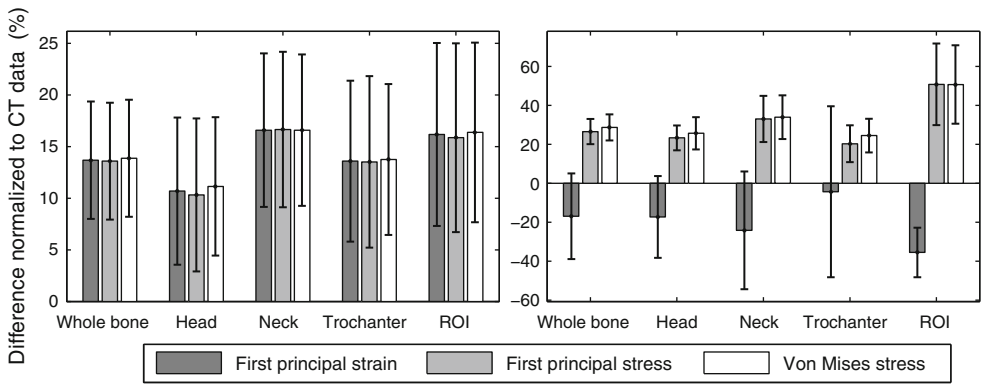

Fig. 7 The difference in mean first principal strain, first principal stress and Von Mises stress between the FE models based on the CT images or the estimated shape. The results (mean ± SD) are presented for different regions of the proximal femur, after normalizing them to the CT data. On the *left side*, the results when using a constant Young's modulus are shown. On the *right side*, the results when using a non-constant (image-based) 2D Young's modulus are shown. With constant Young's modulus, all parameters showed similar errors, whereas with varying Young's modulus the first principal strain of the estimated shape model had 15% lower values, and the fist principal stress and von Mises stresses were 28% lower

tomical region of the proximal femur. The differences for the whole proximal femur were 14 ± 6%. However, when using a non-constant (the BMD image-based) Young's modulus, the normalized values of the first principal strain differed considerably from the normalized values of the first principal stress and von Mises stress. The normalized first

Table 2 Linear Pearson's correlations (r^2) for the first principal strain, first principal stress and von Mises stress between the original CT shape and the estimated shape in six anatomical locations of the femur

	Constant Young's modulus			2D Young's modulus		
	Strain	Stress	Von Mises	Strain	Stress	Von Mises
Whole proximal femur	0.92**	0.93**	0.93**	0.71**	0.91**	0.91**
Head	0.92**	0.92**	0.92**	0.82**	0.97**	0.95**
Neck	0.89**	0.89**	0.89**	0.37	0.78**	0.79**
Trochanter	0.85**	0.84**	0.86**	0.56*	0.67**	0.72**
Shaft	0.92**	0.92**	0.91**	0.61**	0.66*	0.60**
ROI	0.91**	0.91**	0.91**	0.68**	0.70**	0.71**

On left-hand side, the correlations when using a constant Young's modulus value are shown and on the right-hand side, the correlations when using a non-constant (image-based) 2D Young's modulus are shown. With constant Young's modulus, the correlations were all significant and above 0.84. With non-constant Young's modulus, the correlations were mostly significant, although somewhat weaker. In the specific ROI in the neck, the correlations were ranging from 0.68 to 0.71

*$p < 0.05$; **$p < 0.01$

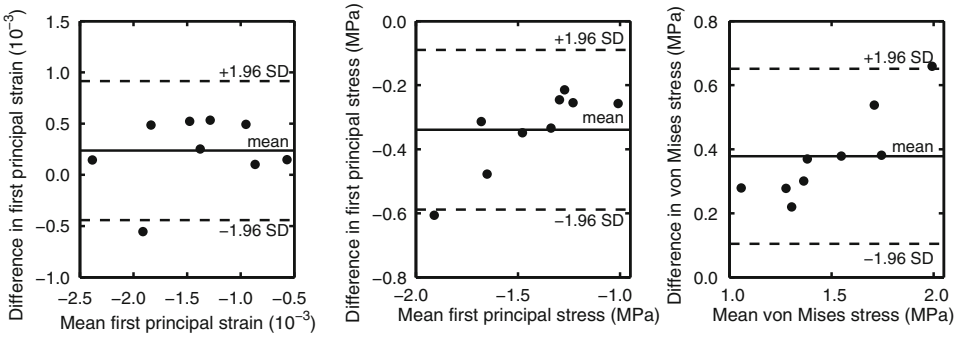

Fig. 8 The Bland–Altman plots of mean first principal strain, first principal stress and von Mises stress for the whole proximal femur when non-constant (image-based) 2D Young's modulus was used (Bland and Altman 1986). The difference between modelling parameters of models based on original CT shape and estimated shape is plotted against the average value of the modelling parameters. The solid line represents the mean difference and the *dashed lines* ±1.96 standard deviations

principal strain values were 17 ± 22% lower than that of the CT-based model whereas the normalized first principal and the von Mises stresses were 27 ± 6 and 29 ± 7% higher than those of the CT model. Similar behaviour was observed in all anatomical regions.

The correlations of the mechanical parameters between the original femur and the estimated shape model were generally significant (Table 2). With a constant Young's modulus, all correlations between the original and estimated shape models were above 0.84 ($p < 0.01$). When a non-constant Young's modulus was used, the correlations were somewhat weaker (Table 2). The correlations for the whole proximal femur remained high for stresses (0.91, $p < 0.01$), whereas the correlations of the first principal strain were lower (neck, $r = 0.37$, non-significant, and trochanter $r = 0.56$, $p < 0.05$).

The Bland–Altman plots indicated that the difference in the first principal strain was independent of the strain value, whereas that in the first principal stress and von Mises stress cases increased systematically as a function of stress (Fig. 8).

4 Discussion

In this study, we applied a method to estimate the bone shape in 3D, using a 2D BMD image and a femur shape template, and assessed the accuracy of the shape estimation method. Further, we conducted a FE analysis of the proximal femur using both the original CT images and the estimated shape to evaluate the differences in values of mechanical parameters. The errors induced by the shape estimation method itself were found to be minor. In the shape estimation method, as expected, the number of bones used to create the template was an important factor for the accuracy of the estimation. Further, small errors in shape estimation were found to induce larger errors in the FE simulations.

4.1 Self-assessment of shape

The errors resulting from the shape estimation method itself were found to be minor by comparing the original and the

re-created proximal femur shape (Table 1). The maximum errors were located close to the major trochanter. This can be explained by the highly irregular shape of the trochanter, and the gaps that exist in the anterior-posterior direction, which the shape estimation method was unable to detect. However, this specific part of the trochanter is unlikely to be the most significant in prediction of fractures induced by the loading conditions used (Keyak et al. 2001). The most important part for loading conditions of this type, the femoral neck (Keyak et al. 2001), resulted in very small errors (Table 1). On the other hand, in hip fractures resulting from a fall on the side, the load is focused on the major trochanter, so in these cases the trochanter area is more important for fracture risk assessment (Keyak et al. 2001). Therefore, the shape estimation method requires some improvement in this region to be able to predict risk for such specific fractures.

4.2 Shape estimation

When using the shape template to predict the geometry of the femurs, the mean volumetric differences and the mean and maximum distance differences indicated that the femoral head and neck were more accurately estimated than the more irregular areas, such as the trochanter. The volumetric difference between the original and the estimated femoral shape was independent of the number of bones used to create the template. However, the mean distance difference between the original and estimated shape decreased significantly from 1.4 to 1.2 mm ($p < 0.01$). In a previous study, Langton et al. (2009b) reported that the mean thickness error between the original and estimated shape decreased from 3.4 to 1.7 mm in the anterior-posterior direction when the pixel size decreased from 1.080 to 0.674 mm. Moreover, Kolta et al. (2005) reported a mean absolute error of 0.8 mm between the estimated and CT shape surfaces with the pixel size of 0.25 mm, using the shape estimation based on two orthogonal DXA images. Hence, based on the present and previous results it seems that the estimation accuracy may be inversely related to the pixel size.

To get as realistic an assessment of the shape estimation accuracy as possible, future comparisons between the original and estimated shapes should be performed in three dimensions, not merely in the anterior-posterior direction (Langton et al. 2009b). Further, in the shape re-creation accuracy analysis, the comparisons should be performed between the shapes, not just between the landmarks used in shape deformation (Langton et al. 2009b). In addition, our findings suggest that analysis at different anatomical parts of the femur may lead to a more accurate validation of the methods when evaluating their ability to predict the femoral shape and mechanical behaviour.

Our results indicate that the choice of the template bones may be critical for accurate shape estimation. The large standard deviations in Fig. 6 highlight the importance of the proper template. A well-chosen template leads to a volumetric difference of only a few percents and a mean distance difference of less than 1.0 mm, whereas in the worst case the errors are 15% and 1.6 mm. This suggests that, ultimately, several different templates should be available, and the template for each patient should be chosen based on several criteria, such as gender, age, size, or the length and width of the femoral neck. In future studies, we hope to include such criteria and evaluate the potential changes in the estimation accuracy.

4.3 Finite element modelling

The errors in the 3D shape of the proximal femur will translate into errors in FE analyses when the model geometry is based on the estimated shape. To determine how big these errors are, we studied the agreement and linear correlation of mechanical parameters, derived from the original CT shape and estimated shape. Then, both the constant Young's modulus value and the value that varied in 2D were applied (Fig. 7). With the constant Young's modulus value, even when the absolute errors were high, the correlations between the original and estimated shapes were significant ($r^2 > 0.84$, $p < 0.01$). This suggests that the errors were mostly systemic but the individual variation in strength could be predicted.

When the 2D spatial variation in Young's modulus was applied, the associations between the parameters values from the CT and estimated shape were somewhat weaker, and the differences in parameter values larger, than with the constant Young's modulus distribution. When a more realistic material model is used, the simulation outcome seems to be more sensitive to shape estimation errors. The trends in Bland–Altman plots (Fig. 8) for the first principal stress and von Mises stress cases indicate that the errors in these parameters were dependent on the value of the parameter. From the modelling parameters used, the stiffness had higher correlations ($r^2 = 0.83$, $p < 0.01$) than other parameters (in ROI region r^2 is between 0.68, $p < 0.01$ and 0.71, $p < 0.01$).

This study applied the most developed methods to estimate the strength of the proximal femur based on one BMD image, 2D varying Young's modulus and simulated stance phase loading (Langton et al. 2009b). However, the method could not quantify the internal 3D architecture or density within the femur and was lacking information on where cortical bone ends and trabecular bone starts. To develop the method further, we aim to apply a three-dimensional estimate for Young's modulus which could model the true femoral strength more accurately. Our future studies will include additional loading cases representing typical falls leading to fractures, a non-linear material model for more realistic

characterization of bone behaviour under loading and a yield criterion for more accurate fracture risk assessment.

In conclusion, the present study demonstrates that the shape estimation method applied, based on one BMD image, can predict the 3D shape of the proximal femur with moderate accuracy. However, in the FE analysis of the proximal femur based on the estimated shape, these relatively small errors in shape may induce larger errors in mechanical parameters. This highlights the need for improvements in the method when aiming for accurate assessment of fracture risk using X-ray images and BMD information.

Acknowledgments The authors acknowledge CSC- the Finnish IT Center for Science for computational tools, and the Academy of Finland (128863), European Commission (BONEQUAL-219980), Kuopio University Hospital (EVO, n:o 2009/00117), and national graduate school of musculoskeletal disorders and biomaterials for funding.

References

Bergmann G, Deuretzbacher G, Heller M et al (2001) Hip contact forces and gait patterns from routine activities. J Biomech 34: 859–871

Bland M, Altman D (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 327: 307–310

Bookstein F (1989) Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans Pattern Anal 11:567–585

Cody D, Gross G, Hou FJ et al (1999) Femoral strength is better predicted by finite element models than QCT and DXA. J Biomech 32:1013–1020

Dryden I, Mardia K (1998) Statistical shape analysis. Wiley, New York

Edwards W, Gillette J, Thomas J et al (2008) Internal femoral forces and moments during running: implications for stress fracture development. Clin Biomech 23:1269–1278

Fang Q, Boas DA (2009) Tetrahedral mesh generation from volumetric binary and grayscale images. In: International symposium on biomedical imaging: from nano to macro. IEEE, Boston, Massachusetts, USA

Filippi S, Motyl B, Bandera C (2007) Analysis of existing methods for 3D modelling of femurs starting from two orthogonal images and development of a script for a commercial software package. Comput Meth Prog Bio 89:76–82

Gower J (1975) Generalized procrustes analysis. Psychometrika 40: 33–51

Keyak J, Falkinstein Y (2003) Comparison of in situ and in vitro CT scan-based finite element model predictions of proximal femoral fracture load. Med Eng Phys 25:781–787

Keyak J, Rossi S, Jones K et al (2001) Prediction of fracture location in the proximal femur using finite element models. Med Eng Phys 23:657–664

Keyak J, Rossi S, Jones K et al (1998) Prediction of femoral fracture load using automated finite element modeling. J Biomech 31: 125–133

Kolta S, Le Bras A, Mitton D et al (2005) Three-dimensional X-ray absorptiometry (3D-XA): a method for reconstruction of human bones using a dual X-ray absorptiometry device. Osteoporos Int 16:969–976

Langton C, Pisharody S, Keyak J (2009) Comparison of 3D finite element analysis derived stiffness and BMD to determine the failure load of the excised proximal femur. Med Eng Phys 31:668–672

Langton C, Pisharody S, Keyak J (2009) Generation of a 3D proximal femur shape from a single projection 2D radiographic image. Osteoporos Int 20:455–461

Laporte S, Skalli W, De Guise J et al (2003) A biplanar reconstruction method based on 2D and 3D contours: application to the distal femur. Comput Methods Biomech Biomed Engin 6:1–6

Morgan E, Bayraktar H, Keaveny T (2003) Trabecular bone modulus–density relationships depend on anatomic site. J Biomech 36: 897–904

Pasco J, Seeman E, Henry M et al (2006) The population burden of fractures originates in women with osteopenia, not osteoporosis. Osteoporos Int 17:1404–1409

Wirtz D, Pandorf T, Portheine F et al (2003) Concept and development of an orthotropic FE model of the proximal femur. J Biomech 36:289–293