LUND UNIVERSITY

Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image

Väänänen, Sami P; Grassi, Lorenzo; Flivik, Gunnar; Jurvelin, Jukka S; Isaksson, Hanna

Published in:
Medical Image Analysis

DOI:
10.1016/j.media.2015.06.001

2015

Link to publication

Citation for published version (APA):
Väänänen, S. P., Grassi, L., Flivik, G., Jurvelin, J. S., & Isaksson, H. (2015). Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image. *Medical Image Analysis*, *24*(1), 125-134. https://doi.org/10.1016/j.media.2015.06.001

Total number of authors:
5

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Accepted Manuscript

Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image

Sami P. Väänänen , Lorenzo Grassi , Gunnar Flivik ,
Jukka S. Jurvelin , Hanna Isaksson

PII: S1361-8415(15)00088-2
DOI: 10.1016/j.media.2015.06.001
Reference: MEDIMA 1012

To appear in: *Medical Image Analysis*

Received date: 3 August 2014
Revised date: 3 June 2015
Accepted date: 11 June 2015

Please cite this article as: Sami P. Väänänen , Lorenzo Grassi , Gunnar Flivik , Jukka S. Jurvelin , Hanna Isaksson , Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image, *Medical Image Analysis* (2015), doi: 10.1016/j.media.2015.06.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Highlights

- 3D shape, vBMD and FE mesh of the proximal femur were reconstructed from a DXA image
- Surface reconstruction error was 1.0-1.4mm
- Use of DXA, instead of projected CT images, had no significant effect on the reconstruction accuracy
- DXA-based FE agreed well with CT-based FE analysis (stiffness r^2=0.85, MAC=0. 0.977)
- DXA-based FE analysis may help to simulate patient-specific bone mechanics

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Generation of 3D shape, density, cortical thickness and finite element mesh of proximal femur from a DXA image

Sami P. Väänänen[a,b,c*], Lorenzo Grassi[d], Gunnar Flivik[e], Jukka S. Jurvelin[a,c], Hanna Isaksson[d,e]

[a] Department of Applied Physics, University of Eastern Finland, Kuopio, Finland, POB 1627, 70211 Kuopio, Finland

[b] Department of Clinical Physiology and Nuclear Medicine, Kuopio University Hospital, Kuopio, Finland, POB 1777, 70211 Kuopio, Finland

[c] Diagnostic Imaging Centre, Kuopio University Hospital, Kuopio, Finland, POB 100, 70029, Kuopio, Finland

[d] Department of Biomedical Engineering, Lund University, Lund, Sweden, Box 118, 221 00 Lund, Sweden

[e] Department of Orthopaedics, Clinical Sciences, Lund University, Lund, Sweden, SE-221 85 Lund, Sweden

EMAIL ADDRESSES:
sami.vaananen@uef.fi, lorenzo.grassi@bme.lth.se, gunnar.flivik@med.lu.se, jukka.jurvelin@uef.fi, hanna.isaksson@bme.lth.se

CORRESPONDING AUTHOR:
Sami Väänänen, M. Sc.
Department of Applied Physics, University of Eastern Finland
POB 1627, FIN-70211, Kuopio, Finland
Phone: +358403552273; Fax: +358 17 162585
E-mail: sami.vaananen@uef.fi

RUNNING TITLE: 3D FE mesh from 2D DXA image

WORD COUNT: Abstract (371), Main text (7086)

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

Areal bone mineral density (aBMD), as measured by dual-energy X-ray absorptiometry (DXA), predicts hip fracture risk only moderately. Simulation of bone mechanics based on DXA imaging of the proximal femur, may help to improve the prediction accuracy. Therefore, we collected three (1-3) image sets, including CT images and DXA images of 34 proximal cadaver femurs (set1, including 30 males, 4 females), 35 clinical patient CT images of the hip (set 2, including 27 males, 8 females) and both CT and DXA images of clinical patients (set 3, including 12 female patients). All CT images were segmented manually and landmarks were placed on both femurs and pelvises. Two separate statistical appearance models (SAMs) were built using the CT images of the femurs and pelvises in sets 1 and 2, respectively. The 3D shape of the femur was reconstructed from the DXA image by matching the SAMs with the DXA images. The orientation and modes of variation of the SAMs were adjusted to minimize the sum of the absolute differences between the projection of the SAMs and a DXA image. The mesh quality and the location of the SAMs with respect to the manually placed control points on the DXA image were used as additional constraints. Then, finite element (FE) models were built from the reconstructed shapes. Mean point-to-surface distance between the reconstructed shape and CT image was 1.0mm for cadaver femurs in set 1 (leave-one-out test) and 1.4mm for clinical subjects in set 3. The reconstructed volumetric BMD showed a mean absolute difference of 140 and 185mg/cm^3 for set 1 and set 3 respectively. The generation of the SAM and the limitation of using only one 2D image were found to be the most significant sources of errors in the shape reconstruction. The noise in the DXA images had only small effect on the accuracy of the shape reconstruction. DXA-based FE simulation was able to explain 85% of the CT-predicted strength of the femur in stance loading. The present method can be used to accurately reconstruct the 3D shape and internal density of the femur from 2D DXA images. This may help to derive new information from clinical DXA images by producing patient-specific FE models for mechanical simulation of femoral bone mechanics.

Highlights

- 3D shape, vBMD, cortical thickness and FE mesh of the proximal femur were reconstructed from a DXA image
- Surface reconstruction error was 1.0-1.4mm
- Use of DXA, instead of projected CT images, had no significant effect on the reconstruction accuracy
- DXA-based FE agreed well with CT-based FE analysis (stiffness r^2=0.85, MAC=0.977)
- DXA-based FE analysis may help to simulate patient-specific bone mechanics

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Keywords

Shape reconstruction, Finite element, Proximal femur, DXA, Bone mineral density, Statistical appearance model

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1. *Introduction*

Osteoporosis is a disease characterized by reduced mineral density and inferior bone quality. Therefore, bone fracture risk increases with osteoporosis (Kanis et al., 1994; WHO, 1994). It is estimated that 30% of all women over 50 years of age in Europe are affected by osteoporosis (Melton et al., 1992). The number of osteoporotic fractures in Europe was estimated in year 2000 to be 3.1-3.7 million, and their direct cost €32 billion (Kanis and Johnell, 2005). These costs are expected to rise to €76.8 billion per year in 2050 (Kanis and Johnell, 2005).

Bone mineral density (BMD) measurement using dual-energy x-ray absorptiometry (DXA) at the hip is most commonly used to diagnose osteoporosis (Link, 2012). The patient's BMD value is compared with the mean BMD value of the reference population of healthy young women. The result is reported as a T-score, which is the number of standard deviations (SD) from the mean population (WHO, 1994). A patient is defined as osteoporotic if the T-score value is below -2.5 SD. However, BMD alone can explain less than one third of all low trauma fractures (Pasco et al., 2006). This is partly due to facts that BMD does not account for the geometry and internal architecture of the bone. Hip structural analysis (HSA) measures geometric features of the femoral shape e.g. femoral neck length and cross-sectional moment of inertia from a DXA image. Although HSA structural parameters are highly correlated with BMD, they are not better predictors of fracture risk than BMD (Bonnick, 2007). Finite element (FE) modelling based on quantitative CT (QCT) images has been shown to explain at least 20% more of the strength of the femur (Cody et al., 1999). With QCT-based FE modelling, additional factors that help explaining femoral strength can be taken into account. These include e.g. shape and structure of the trabecular bone (Larsson et al., 2014), the effect of the loading direction (Grassi et al., 2012; Trabelsi and Yosibash, 2011) and the thickness of the cortical bone (Chevalier et al., 2009). It has also enabled prediction of the fracture location (Schileo et al., 2014). However, CT imaging is more expensive and induces larger radiation dose on the patients than DXA (Kanis, 2002). This impairs the use of CT-based FE models for fracture risk prediction in the clinics.

Two-dimensional FE models can be built from DXA images (Luo et al., 2013; Naylor et al., 2013; Sarkalkan et al., 2014). Femoral strength computed with 2D FE models was shown to be an independent

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

predictor of fracture risk (Naylor et al., 2013). However, 2D FE analysis was also shown to be sensitive to the patient's position (Luo et al., 2013). It is also intrinsically not able to account for the impact in the anterior-posterior direction, which is known to significantly affect femur's load-bearing ability (Ford et al., 1996). DXA-based 3D FE analyses could overcome these limitations.

If the shape and internal density of a bone could be reconstructed from one or a few 2D radiographs, it would provide information about the 3D bone geometry and architecture for the assessment of the fracture risk. Thus, the concept of shape reconstruction has been studied intensively (Ahmad et al., 2010; Chernoff and Nielsen, 2009; Galibarov et al., 2010; Iwashita et al., 2007; Knaan and Joskowicz, 2003; Kurazume et al., 2009; Langton et al., 2009; Markelj et al., 2012; Whitmarsh et al., 2013, 2011a; Zheng, 2009). Most of the previous methods use several 2D images, and only a few studies reconstructed both shape and internal density of the proximal femur using one 2D radiograph (Galibarov et al., 2010; Thevenot et al., 2013; Whitmarsh et al., 2011; Väänänen et al., 2012b). Two of them used a single 2D DXA image (Whitmarsh et al., 2011; Väänänen et al., 2012c). However, these methods do not automatically yield the 3D mesh for the FE simulation. Further, the methods do not distinguish between trabecular and cortical bone (Thevenot et al., 2013). This is a limitation since the geometry and thickness of the cortical bone play an essential role in determining the strength of the total femur (Holzer et al., 2009; Koivumäki et al., 2012a). Therefore, an automatic method that accurately estimates bone 3D shape and internal density and generates FE meshes from a single clinical 2D DXA image is still missing.

The aim of the present study was to develop an automatic method to reconstruct the 3D shape, internal volumetric BMD (vBMD) and cortical thickness of the proximal femur from a single hip DXA image. Further, the outcome should be directly suitable for FE simulations. The method could subsequently be used to calculate bone strength and predict patient's susceptibility to fractures. We assessed the major sources of error and how the errors accumulated during each step of the shape reconstruction. Especially, we evaluated the effect of noise and spatial resolution in DXA imaging on the accuracy of the reconstruction by comparing it to the reconstruction from a projected CT image. We also assessed the suitability of the generated FE-mesh for mechanical evaluation. Finally, the effect of shape reconstruction on the results of the FE models was assessed.

ACCEPTED MANUSCRIPT

2. Material and Methods

2.1. Bone materials

Three different collections of bone images (sets 1-3) were used in this study. The two first sets served as teaching sets for the statistical appearance models, whereas the third set was used to independently test the accuracy of the method.

In set 1, 34 unpaired human proximal femurs (13 right and 21 left; 30 men and 4 women; age range 18-82 years, 50±16 years (mean±SD)) were collected. Ethical permission for the study was granted by the National Authority for Medicolegal Affairs (TEO, 5783/04/044/07). None of the cadavers had any pre-existing conditions that might affect bone metabolism. Each femur was imaged with a clinical CT scanner (Siemens Definition AS64, Siemens AG, Erlangen, Germany) using in-plane resolution of 0.4 mm with 0.6 mm slice spacing, tube voltage 120kV and tube current 336mA. The water-free dipotassium phosphate (K_2HPO_4) content of each voxel was determined using a CT calibration phantom (Mindways Software Inc., Austin, TX, USA). All bones were also imaged with Lunar Prodigy DXA scanner (GE Healthcare, Madison, WI) at 1.05x0.60 mm resolution. The mean areal BMD measured with Lunar Prodigy at the neck was 0.893±0.164 g/cm^2. In addition, 11 of the bones were imaged with Lunar iDXA DXA scanner (GE Healthcare, Madison, WI) at a resolution of 0.25x0.30 mm.

In set 2, pre-operative CT scans (Philips Brilliance 64, Philips Healthcare, The Netherlands, in-plane resolution between 0.7-1.0 mm, slice separation 0.5 mm, tube voltage 140 kV and tube current 60 mA) of 35 patients undergoing total hip arthroplasty (27 men and 8 women, age range 39-74, 58.1 ± 8.7 years) were collected. Ethical permission was granted by The Ethics Committee at Lund University, Faculty of Medicine (permission 2009/369). A different phantom (Scandinavian Customized Prosthesis) was used when scanning set 2. The phantom was cross-calibrated against the phantom used for set 1 using same CT parameters as for set 2, and equivalent K2HPO4 values were obtained for each voxel in set 2. In this study, images of the hips on the contralateral side from the hip undergoing arthroplasty were used.

In set 3, the left hips of 14 patients (all females, age range 69-78, 74.5 ± 2.7 years) were imaged with CT (Philips Precedence 6P, in-plane resolution between 0.7-0.8 mm, slice separation 2.0 mm, tube voltage 120

kV, tube current 100 mAs) and DXA (GE Lunar Prodigy) (Karjalainen et al., 2012). Five of the patients had previous fracture in the femur, two at the left side. The two patients who had fracture at the left side were excluded from the evaluation since they had intramedullary nails in their femur. The same Mindways BMD phantom as in set 1 was used in these CT images. The ethical permission was granted by the local ethical committee (Kuopio University Hospital Ethical Committee, permission 80/2008).

All CT images were pre-processed using Mimics (version 15.1, Materialize, Belgium). First, the CT images in set 1 were resampled to an isotropic voxel size of 0.4 mm, whereas a voxel size of 0.7 mm was used for set 2 and 3. All bones were manually segmented from the surrounding tissues using thresholding and basic morphological operations. Possible defects were corrected manually. Point clouds, which were used later in the method, were created from the segmented bone surfaces. A triangular surface mesh was first generated for the bone surfaces using the isosurface function in Matlab (release R2012a, MathWorks, Inc., Natick, Massachusetts, United States). The remaining stair step artefacts in the surface mesh were thereafter smoothed with a low-pass filter (Bade, Haase, & Preim, 2006) implemented in the iso2mesh toolbox v1.5.0 (Fang & Boas, 2009). The mesh vertices together with the centres of the surface triangles formed the point cloud of the surface. For the femur, the endocortical bone surface, i.e. the interface between the cortical and trabecular bone, was also segmented. First, the minimum distance between the endocortical and periosteal bone surfaces was defined for the cadaver femurs in set 1 by eroding the segmented femur with a mask which radius was 2 voxels (0.8 mm) and for the femurs in set 3 with a mask which radius was 1 voxel (0.7 mm) (Koivumäki et al., 2012b). Then, the rest of the endocortical bone surface was set by distinguishing the cortex from the trabecular bone using thresholding and morphological operations. Again, possible defects were corrected manually. A point cloud was generated for the endocortical bone surface similarly as for the periosteal bone surface.

2.2. Statistical appearance model (SAM)

A SAM describes the variation of shape and density (in the population that is used to train the model) with a few parameters called modes. It is based on reduction of dimensionality of the shape model with principal component analysis (Cootes et al., 2001). Two separate SAMs, one for the femur and one for the

ACCEPTED MANUSCRIPT

pelvis, were built in this study. The following steps were taken when the SAMs were trained based on the CT images of the proximal cadaver femurs in set 1 or the pelvises in set 2 (Fig. 1) (Väänänen et al., 2012a):

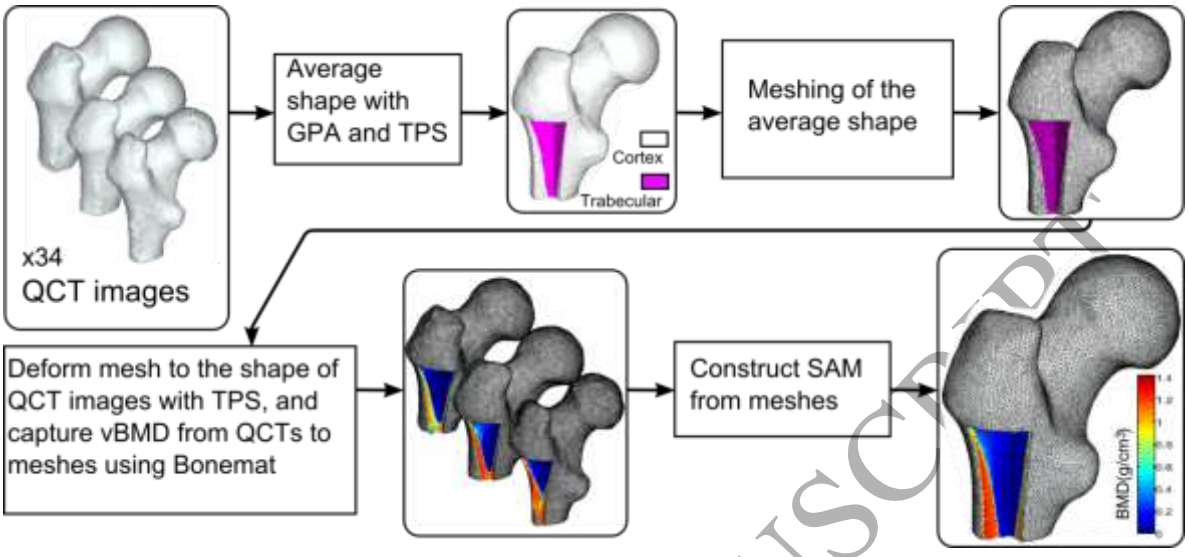


Figure 1: Development of the statistical appearance model (SAM). An average shape was created from QCT images using generalized Procrustes analysis (GPA) and thin-plate splines (TPS), and a mesh was created from the shape. The cortex and trabecular bone were separated in the mesh. The mesh was deformed to the shape of each QCT and BMD was captured for each element of the mesh. The SAM was constructed from the deformed mesh.

1. To create a SAM, a set of anatomical landmarks were placed on the segmented training bones (Väänänen et al., 2012b, 2011). For the femur, 28 landmarks were set automatically. For the pelvis, 20 landmarks were set manually. The landmark positions are described in more detail in the supplementary material.

2. An average template bone was generated (Väänänen et al., 2012b):

    a. Rigid transformations and scaling were removed between anatomical landmarks of the training bones with generalized Procrustes analysis (GPA) (Gower, 1975). Thereafter, the average of each anatomical landmark was calculated.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

b. CT images of the training bones were warped to the shape described by the average landmarks with thin-plate splines (TPS) (Bookstein, 1989).

c. The template image was created by defining a voxel as bone if it was covered by over half of the training bones. For the femoral template image, the bone voxels were determined as trabecular or cortical bone depending which one was the more prevalent one among the covering training voxels. The minimum cortical thickness was set to 0.8 mm.

3. A template mesh was created by meshing the template image using Hypermesh (11.0, Altair Engineering, Inc., Troy, USA) with four node tetrahedral elements. The mesh of the femur had 283 000 nodes and 1 637 000 elements (maximum edge length of 1.5 mm), whereas the mesh of the pelvis had 84 893 nodes and 451 229 elements (maximum edge length of 2.0 mm). These two meshes constructed the SAMs. In the mesh of the femur, the endocortical bone surface was preserved and each element was defined as trabecular or cortical.

4. The template mesh was warped to the shape of each training bone:

a. In an initial warp, the average mesh was warped with TPS to the shape of each training bone based on the landmarks.

b. Total of 2000 surface nodes were randomly chosen from the periosteal and from the endocortical bone surfaces of the mesh. The nearest neighbours from these points were then identified accordingly from the point clouds at the periosteal and endocortical bone surfaces of the training bone. In the pelvis, however, the endocortical bone surface was not used.

c. The average mesh was warped again more accurately to the shape of the training bone with TPS based on the 2000 surface points.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

d. Rigid transformations and scaling between the average and the training meshes were removed with GPA. The average of the scaling factors of the bones $\overline{s_{\mathrm{GPA}}}$ was defined for later use. Thereafter, for each subject, vector $\boldsymbol{x}_{\mathrm{GPA}}$ was formed which included all k node coordinates $\boldsymbol{x}_{\mathrm{GPA}} = (x_1,y_1,z_1,\ldots, x_k,y_k,z_k)^T$ of the mesh in one column vector. ^{T} means matrix transpose.

5. BMD values were captured for each element of the warped mesh using the Bonemat V2 algorithm (Taddei et al., 2007; Viceconti et al., 2007). The algorithm uses Gaussian quadrature to integrate numerically over the CT images at the volume of each element. Before integration, Hounsfield units were converted to BMD using a calibration curve from the BMD phantom. During integration, the BMD distribution was estimated with a piece-wise linear function. The BMD values of the mesh for each subject ($\boldsymbol{g}$) were normalized as $g_{\mathrm{normalized}} = (\boldsymbol{g} - \overline{\boldsymbol{g}})/\boldsymbol{g}_{SD}$, where $\overline{\boldsymbol{g}}$ stands for the mean value and $\boldsymbol{g}_{SD}$ for standard deviation. The average of the $\boldsymbol{g}_{SD}$ (scaling) of the bones $\overline{s_{\mathrm{BMD}}}$ was calculated for later use.

6. The shape and density information of all warped meshes were collected into one matrix $\boldsymbol{X}$. Column j of the matrix $\boldsymbol{X}$ included the node coordinates and the BMD values of each element of the warped mesh for bone j:$\boldsymbol{X}_j = [\boldsymbol{x}_{\mathrm{GPA},j}{}^T, \boldsymbol{g}_{\mathrm{normalized},j}{}^T]^T$. Variation in the coordinates and the BMD values must be uniform. Therefore each row i of the matrix was normalized separately: $\boldsymbol{X}_{\mathrm{normalized},i} = (\boldsymbol{X}_i - \bar{X}_i)/X_{SD,i})$. In this notation $\bar{X}_i$ is the mean and $X_{SD,i}$ standard deviation of row i, and they are collected for later use into column vector $\overline{\boldsymbol{X}}$ and diagonal matrix $\boldsymbol{W}$, respectively.

7. The principal modes of variations, i.e. eigenvectors ($\boldsymbol{\phi}$), and their variances, i.e. eigenvalues (λ), were calculated for the matrix $\boldsymbol{X}_{\mathrm{normalized}}$ with truncated singular value decomposition. Therefore, the shape and the density of bone j can be presented with a linear combination of the eigenvectors:

$$\boldsymbol{X}_j = \overline{\boldsymbol{X}} + \boldsymbol{W}\left(\sum_{m=1}^{c} \boldsymbol{\phi}_m \lambda_{m,j}\right), \tag{1}$$

where c is the number of bones. Significance of an eigenvector is determined as $S_{\boldsymbol{\phi_m}} = \left(\lambda_{\mathrm{SD},m}\right)^2 / \sum_{i=1}^{c} \lambda_{\mathrm{SD},i}{}^2$, where $\lambda_{\mathrm{SD},m}$ is the SD of the m:th eigenvalue of the training bones. The shape and density distribution of a bone outside the training set can be approximated with a linear combination of the c' most significant eigenvectors

$$\widehat{\boldsymbol{X}} = \overline{\boldsymbol{X}} + \boldsymbol{W}\left(\sum_{m=1}^{c'} \boldsymbol{\phi}_m b_m\right), \text{where} \tag{2}$$
$$\widehat{\boldsymbol{X}} = \begin{bmatrix} \widehat{\boldsymbol{x}}_{\mathrm{GPA}} \\ \widehat{\boldsymbol{g}}_{\mathrm{normalized}} \end{bmatrix}$$

and $\boldsymbol{b}$ includes the eigenvalues, i.e modes, of the bone outside the training set. Finally, the approximated bone needs to be scaled to the original average size and BMD scale: $\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{x}}_{\mathrm{GPA}} \overline{s_{\mathrm{GPA}}}$ and $\widehat{\boldsymbol{g}} = \widehat{\boldsymbol{g}}_{\mathrm{normalized}}\, \overline{s_{\mathrm{BMD}}}$.

2.3. Reference images

Five different reference images were used in the reconstruction of the 3D femoral shape: 2D projection of the CT image of the cadaver femurs along the anterior-posterior plane (set 1), images from the iDXA and Prodigy DXA scanners of the cadaver femurs (set 1), and CT projection and image from Prodigy DXA scanner of the clinical subjects (set 3) (Fig. 2). The DXA images were segmented with a threshold of 160 mg/cm^2 to determine the region covered by bone (B_{REF}). Thereafter the largest connected object was chosen and image closing was performed using a mask with radius of 1mm. Landmarks were positioned on each reference image. Landmarks 1 and 2 were placed at the intersections of the contours of the shaft and the minor trochanter, and landmarks 3-5 were set into the joint space such that they formed roughly a half circle. Four extra landmarks were set on the clinical reference images to the medial side of the pelvis (Fig. 2).

ACCEPTED MANUSCRIPT

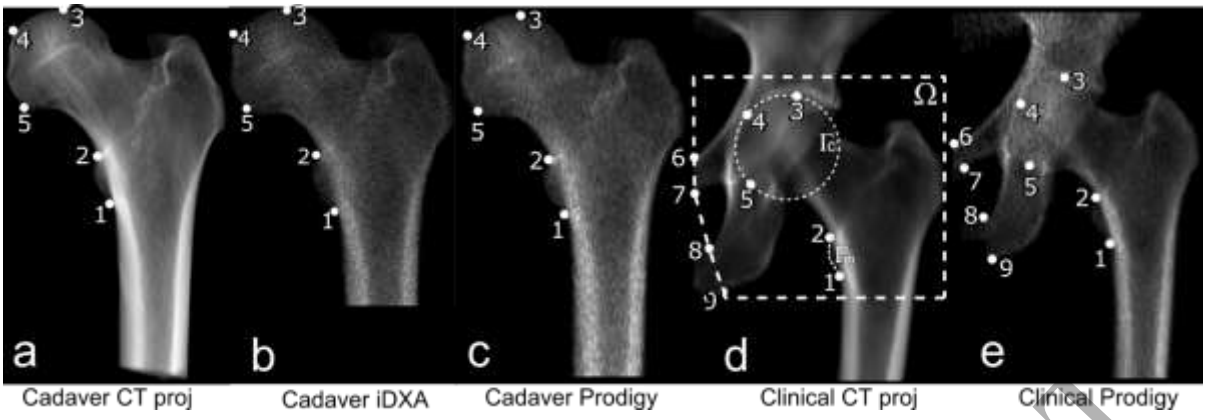


Figure 2: a-c: Reference images, i.e. projected CT image and DXA images obtained with iDXA and Prodigy scanners, of the cadaver femurs from set 1. d-e: Reference images, i.e. CT image and image from Prodigy scanner, of the femur and pelvis from the clinical subjects in set 3. Landmarks were positioned on each reference image. Dashed polygon in sub-figure d bounds the region Ω where the cost function of the registration was calculated. The dashed arc Γ_m in sub-figure d shows the 2D contour of the minor trochanter between landmarks 1 and 2, and circle Γ_c was fitted on the landmarks 3-5. They were used in one part of the calculation of the cost function.

2.4. Reconstruction of shape

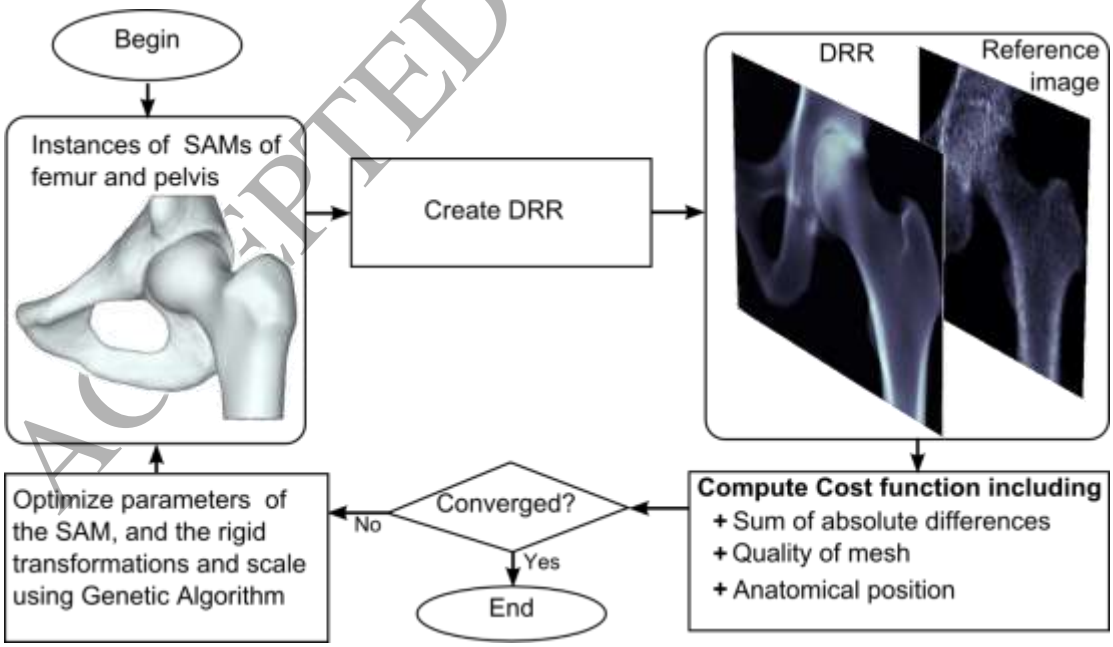


Figure 3: Flowchart describing the iterative reconstruction of 3D shape based on a 2D reference image, which was a projected CT or DXA image.

ACCEPTED MANUSCRIPT

The shape and the internal density of a femur were reconstructed by registering digitally reconstructed radiograph (DRR) (see section 2.5.1) over the 2D reference image (Fig. 3). The genetic algorithm (GA) in Matlab (global optimization toolbox) was used to register and optimize the translations, rotations, scale and the mode values of the SAMs by minimising a cost function. The cost function (C) was defined as the sum of three measures: sum of absolute difference (SAD, C_{SAD}) between the aBMDs of the images (see section 2.5.2), quality of mesh $C_{\mathrm{mesh_qual}}$) (see section 2.5.3) and anatomical positioning ($C_{\mathrm{anat_pos}}$) calculated based on the landmarks (see section 2.5.4)

$$C = C_{\mathrm{SAD}} + C_{\mathrm{mesh_qual}} + C_{\mathrm{anat_pos}} \tag{3}$$

In the initial step, the mode values of the SAMs were set to 0, i.e., the registration began from the average shape. The initial translation of the instance of femoral and pelvis 3D SAMs were calculated by minimizing the measure of the anatomical positioning. The population size in the GA was set to 100 when only the femur was registered (set 1) and set to 200 when both the femur and the pelvis were registered (set 3). The GA converged when the cost function did not decrease during 4 (set 1) or 6 (set 3) consecutive generations. Seventeen of the most significant mode values were used in both SAMs. In the femoral SAM these were able to explain 95% of the variation in the training set, whereas in the pelvis SAM they explained 80% of the total variation. The GA needed limits for the optimized parameters. The mode values were allowed to vary ±3 standard deviations, the rotations varied ±10° and the scale varied from 0.8 to 1.2.

2.4.1. Digitally reconstructed radiograph

A DRR was created by projecting an instance of SAM onto the coronal plane. It started with finding the centre of each element, calculated as the average of the coordinates of its vertices. The BMD value of the element was granted for this centre point. Then, a density value was given for each vertex by calculating the average density value of the elements which shared the vertex. The mesh was thereafter converted into a 3D image. In the image, the voxel size in the coronal plane was the same as the pixel size in the reference image. The voxel size in the third direction was the average of the other two ones. The 3D image enclosed the instance of the 3D SAM and a density value was given for all voxels which enclosed a vertex or a centre of

ACCEPTED MANUSCRIPT

an element. The created 3D image was sparse, *i.e.*, some voxels inside the mesh remained without a BMD value. These voxels were detected by closing the image using the smallest cubic mask that filled the sparse image. A density value was given for each new object voxel from the nearest vertex or centre of an element using the nearest neighbour search. Finally, the created 3D image was projected to 2D by summing the intensities in the anterior-posterior direction. The region that the bones cover in the 2D image was defined as B_{DRR}.

2.4.2. Sum of absolute difference

The cost bearing the density differences between a DRR and a reference image was calculated as the SAD between the areal BMDs of the images. The region **Ω,** where the SAD was calculated, was defined as the region with inferior boundary 5mm below landmark 1, superior boundary 5 mm above landmark 3, and medial boundary passing through landmarks 6-9 in the reference image (Fig. 2d). Finally, the SAD was divided by the number of pixels (*N*) covered by any of the bones in the region Ω, i.e., $(B_{\mathrm{REF}} \cup B_{\mathrm{DRR}})$

$$C_{\mathrm{SAD}} = \frac{1}{N}\sum_{i=1}^{N} |REF - DRR|. \tag{4}$$

Therefore, C_{SAD} ranged from zero to approximately the average areal BMD (~0.9 g/cm^2).

2.4.3. Quality of mesh

The quality of the mesh of the instance of SAM was calculated for each element with scaled minimum solid angle (θ_{min}) (Liu and Joe, 1994). It provides a measure for mesh quality between 0 (poor) and 1 (optimal). Iso2mesh Matlab toolbox (Fang and Boas, 2009) was used to calculate this metric. The cost was defined as

$$C_{\mathrm{mesh_qual}} = \frac{N_\theta}{NE} \tag{5}$$

where N_θ is the number of the elements having $\theta_{\mathrm{min}} < 0.5$, and *NE* is the total number of elements.

2.4.4. Anatomical positioning calculated based on landmarks

ACCEPTED MANUSCRIPT

The cost of anatomical positioning was calculated using landmarks 1-5 at the reference images. First, the 2D contour of the minor trochanter in the reference image between landmarks 1 and 2 (Γ_{m}) was determined (Fig. 2e). Next, the tip of the minor trochanter in the DRR ($DRR_{\mathrm{mintroc_tip}}$) was determined using landmark 4 in the 3D femoral SAM. Landmark 4 was first projected on the DRR. The projected landmark was then moved to the 2D contour of the minor trochanter in the DRR along a line perpendicular to the shaft axis. The minimum distance from the 2D contour in the reference image to the tip of the minor trochanter in the DRR was calculated as $d_{\mathrm{mintroc}} = \min\left(\left\|\Gamma_{\mathrm{m}} - DRR_{\mathrm{mintroc_tip}}\right\|\right)$. The contour of the femoral head in DRR (Γ_{h}) was defined by fitting a ball to the medial part of the head of the femoral SAM. The minimum distance from the femoral head contour to landmarks 3-5 in the reference image was calculated as $d_{\mathrm{min_head}\,i} = \min(\|REF_{\mathrm{landmark}\,i} - \Gamma_{\mathrm{h}}\|)$ and the maximum of them was used: $d_{\mathrm{f_head}} = \max\left(d_{\mathrm{min_head}\,i}\right)$, where i=3,4,5. Anatomical positioning of the pelvis was calculated by first fitting a circle to landmarks 3-5 in the reference image (Γ_{c}) (Fig. 2e). Then, landmarks 1, 2 and 6 of the pelvic SAM model were projected onto the DRR of the pelvis, and the minimum distances from Γ_{c} to the locations of the projected landmarks 1, 2 and 6 were calculated $d_{\mathrm{min_pelvis},i} = \min\left(\left\|\Gamma_{\mathrm{c}} - DRR_{\mathrm{pelvis_landmark},i}\right\|\right)$. The maximum of these three distances was used $d_{\mathrm{pelvis}} = \max\left(d_{\mathrm{min_pelvis},i}\right)$, where i=1,2,6. The cost $C_{\mathrm{anat_pos}}$ was the sum of these distances in millimeters with tolerance of one 1 mm. Multiplication by 0.05 was done to ensure that the $C_{\mathrm{anat_pos}}$, similarly as the other costs, was between 0 and 1 during optimization

$$C_{\mathrm{anat_pos}} = \begin{cases} 0 & d < 1 \\ 0.05(d-1) & d \geq 1 \end{cases}, \text{where} \quad d = d_{\mathrm{mintroc}} + d_{\mathrm{f_head}} + d_{\mathrm{pelvis}} \tag{6}$$

2.5. Evaluation

The errors arising from the creation of the SAM and from reducing the number of used modes to the 17 most significant one were evaluated. These correct mode values were captured from the principal component analysis. The original segmented CT images were compared with the respective instance of the SAM.. In the following, this will be referred to as “SAM correct mode values”. The accuracy of the reconstruction was

ACCEPTED MANUSCRIPT

evaluated with three different combinations of bone sets (table 1). In the first test case, the ability of the method to recreate the shape and the internal density of a femur based on one 2D image was evaluated by registering the femoral SAM trained with cadaver bones from set 1 to the reference images from the same set. In the second case, the method's ability to reconstruct the shape of a femur was evaluated by removing the cadaver femur used for testing from the training set of SAM (leave-one-out test). The SAM registration was performed similarly as in case 1. In both cases, the registration was evaluated with all the available reference images, i.e., with 34 projected CT images and DXA images from the Prodigy device and with 11 DXA images from the iDXA device. In the third test, the two separate SAMs, the femoral SAM (built from set 1) and the SAM of the pelvis (built from set 2), were registered to the reference images in set 3. i.e., DXA image from Prodigy and projected CT images of clinical subjects (including both femur and pelvis). We recall that the SAMs of the femur and pelvis were generated separately with their own mode values and independent positions.

Table 1 Bone sets which were used for training the SAMs, evaluation of the shape reconstruction, and reference images used by reconstruction. The number of samples in sets 1, 2 and 3 were 34, 35 and 12, respectively.

Training bones for SAM	Test set	Reference images in test set
Cadaver femurs, set1 (one femoral SAM)	Cadaver femurs, set 1	CT projection, Prodigy, iDXA
Cadaver femurs, set 1 (leave-one-out, in total 34 femoral SAMs)	Cadaver femurs, set 1	CT projection, Prodigy, iDXA
One SAM from cadaver femurs set 1 and one SAM from pelvises set 2	Clinical hips, set 3	CT projection, Prodigy

2.5.1. Reconstruction accuracy of the shape and internal density

In all cases, the accuracy of the reconstruction was evaluated by comparing the reconstructed shape and the internal density to those of the CT image of the same femur. First, the periosteal surface of the reconstructed femoral shape was co-registered to the periosteal surface of the same femur in the segmented CT image by using an iterative closest point algorithm. Then, the mean and maximum point-to-surface distances between the outer bone surface of the reconstructed shape and the segmented CT images were calculated. An image showing the inter-patient mean point-to-surface distance at each location was drawn for DXA image-based reconstruction of the clinical subjects from set 3 (Fig. 5).

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

The template mesh (from section 2.2) was warped directly over the segmented original CT images in the test set (see the following section). Thereafter, the ratio of the volumetric difference between the original and reconstructed shapes was calculated by summing the volumes of the tetrahedra in the meshes of each shape and by calculating their difference. The mean difference in the reconstructed vBMD and original vBMD derived from the CT images was compared element-by-element, and representative examples (Fig. 6) were drawn. The cortex thickness was calculated from the warped and from the reconstructed meshes as the distance from each triangle centre of the periosteal bone surface to the nearest vertex or triangle centre on the endocortical surface. All these parameters were calculated for the whole bone as well as for three anatomical regions; femoral head, neck, and trochanter (Väänänen et al., 2012b). The mesh quality, which was defined with the scaled minimum solid angle, was also calculated for each reconstructed shape.

2.5.2. Effect of reconstruction accuracy on the FE simulations

To evaluate how the reconstruction of the shape and the internal density affects the results of the FE simulation, FE models were built (Grassi et al., 2012) by warping the template mesh directly over the segmented original CT images. After warping, inhomogeneous elastic material properties were mapped using Bonemat V3 algorithm (Taddei et al., 2007). This version of Bonemat is similar to the one described earlier (Bonemat V2), except that the BMD values are now converted to Young's modulus values before integration over the elements. The density-elasticity relationship $E = 8920(\rho_{\mathrm{K2HPO4}})^{1.83}$ from Morgan and co-workers (2003) was used. A similar approach provided results comparable with the state of the art literature (Grassi et al., 2011). Consequently, the obtained FE models (referred to as "warped model") based on original CT images were used as a reference to evaluate the potential discrepancy in the FE results due to errors in the reconstruction of shape and internal density (referred to as "DXA-based model"). The warped and the reference image-based models (CT projection or DXA-based) were both loaded with a force equal to 1.5 times the bodyweight. The force was applied on the femoral head 10° medially from the femoral shaft axis direction. The force was distributed evenly on the nodes at the head which were distinguished by setting a plane perpendicular to the force vector 5 mm inside the surface of the head (Väänänen et al., 2011). Nodes 2 cm below the minor trochanter were constrained.

ACCEPTED MANUSCRIPT

Although the element quality was globally good in all meshes, in some meshes a few elements (4±6, median 0, maximum 29) experienced severe distortion. These elements cannot be handled by the FE solvers. Therefore, a custom Matlab algorithm was iteratively used to fix the distorted elements:

1. Ansys was called to import the mesh and check for excessively distorted elements (using the CHECK command)
2. If a zero-volume element was detected, one of the nodes was moved by a distance equal to 1/3 of its distance to the nearest adjacent node not belonging to the treated element. The node to be moved and its direction were defined as:
    a. If two coincident nodes were found, one of the two was moved towards the direction normal to the largest element face of the adjacent elements.
    b. If one node was lying on the face of the element formed by the other 3 nodes, the node was moved along the negative normal of the face.

The degree of consistency between the warped and DXA-based models was finally assessed in terms of the predicted nodal displacements. First, the total stiffness was compared between the models. Then, the modal assurance criterion (MAC) (Allemang and Brown, 1982; Allemang, 2003; ANSYS, 2012) between the DXA-based and warped FE models was calculated to provide an indication of the overall correspondence between the models. The MAC compares two sets of real nodal solutions, and the output is a correlation-like scalar between 0 (no correspondence between the result sets) and 1 (perfect correspondence).

3. *Results*

The method was able to reconstruct the shape based on the DXA image of a clinical subject with an average accuracy of 1.42 mm (Table 2, Fig. 4 and 5) and with an average volumetric difference with respect to the original shape of 7%. Use of different reference images, i.e., DXA image or projected CT image, did not affect the accuracy markedly. For example, the average surface reconstruction accuracy changed with only 0.00-0.08 mm. The 3D distribution of the vBMD, based on one DXA image from clinical subjects, was reconstructed with an average difference of 0.19g/cm^3 (Table 2, Fig. 6, supplementary figure S3). The

cortical thickness was estimated with an average accuracy of 0.37mm (Table 2, Fig. 7, supplementary figure S3). The produced FE meshes were suitable for FE simulations, since over 99 % of the elements had an element quality higher than 0.5 as measured with the scaled minimum solid angle.

Table 2: Accumulation of the difference between the reconstructed shape and the original CT images. The difference is evaluated with 5 different shape parameters in the whole bone and in 3 different anatomic regions for all test cases.

Original CT versus reconstructed shape	Femoral part	SAM correct mode values	Recreation			Leave -one-out			Clinical	
			CT proj	iDXA	Prodigy	CT proj	iDXA	Prodigy	CT proj	Prodigy
Mean point-to-surface distance (mm)	Whole bone	0.47	0.87	0.80	0.90	1.02	1.07	1.06	1.33	1.41
	Head	0.42	0.63	0.57	0.70	0.74	0.90	0.81	0.99	1.29
	Neck	0.44	0.77	0.77	0.78	0.95	1.01	0.94	1.12	1.14
	Trochanter	0.51	1.04	0.93	1.06	1.22	1.19	1.25	1.58	1.56
Maximum point-to-surface distance (mm)	Whole bone	2.76	4.64	4.15	4.51	5.24	5.46	5.17	6.28	6.50
	Head	2.01	2.84	2.97	2.92	3.40	3.85	3.48	3.58	3.80
	Neck	1.93	3.15	2.85	3.09	3.67	3.84	3.78	4.00	3.73
	Trochanter	2.60	4.42	4.02	4.40	5.00	5.18	5.04	5.76	5.94
Volumetric difference $\left(\frac{\text{original}-\text{reconstructed}}{\text{original}}\right)$	Whole bone	0.01	0.05	0.04	0.05	0.05	0.06	0.05	0.07	0.07
	Head	0.03	0.05	0.04	0.06	0.07	0.09	0.08	0.08	0.16
	Neck	0.02	0.05	0.05	0.05	0.06	0.06	0.05	0.08	0.09
	Trochanter	0.02	0.10	0.08	0.10	0.09	0.08	0.10	0.14	0.14
Mean difference between vBMDs (g/cm3)	Whole bone	0.07	0.14	0.16	0.14	0.18	0.21	0.18	0.20	0.19
	Head	0.08	0.09	0.09	0.09	0.10	0.11	0.10	0.10	0.10
	Neck	0.09	0.11	0.11	0.11	0.13	0.13	0.13	0.15	0.15
	Trochanter	0.07	0.14	0.15	0.15	0.19	0.20	0.19	0.23	0.20
Mean difference between cortical thicknesses (mm)	Whole bone	0.12	0.19	0.19	0.20	0.25	0.26	0.26	0.35	0.37
	Head	0.05	0.08	0.07	0.08	0.10	0.09	0.10	0.11	0.11
	Neck	0.15	0.22	0.22	0.23	0.29	0.28	0.30	0.38	0.44
	Trochanter	0.14	0.21	0.21	0.21	0.27	0.30	0.28	0.41	0.43

ACCEPTED MANUSCRIPT

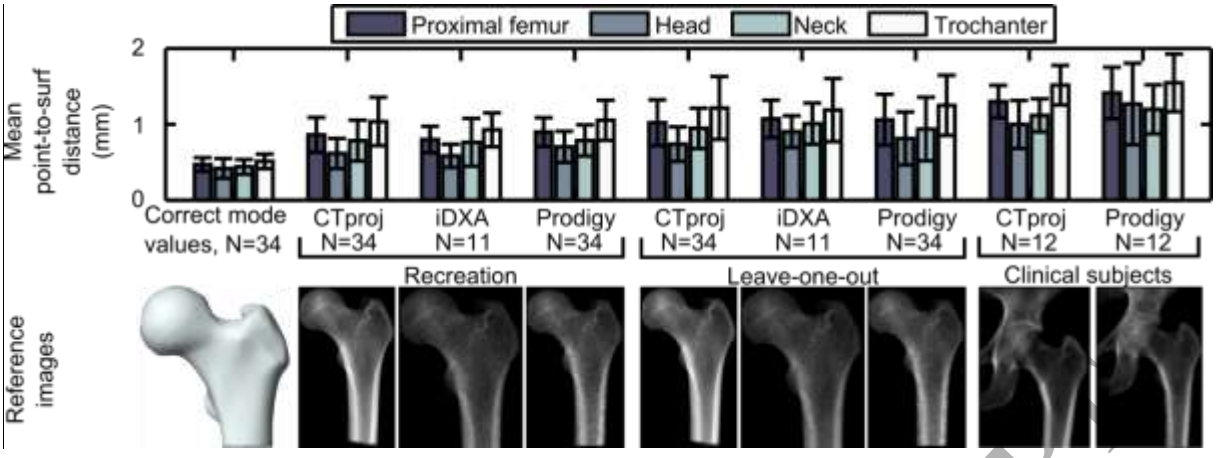


Figure 4: Accumulation of the error during reconstruction. The largest source of error were the creation of the SAM, reconstruction based on only one 2D image and use of clinical DXA images with non-optimal training set. The head and the neck were reconstructed more accurately than the other anatomical regions. Other parameters are presented in supplementary figure S3.

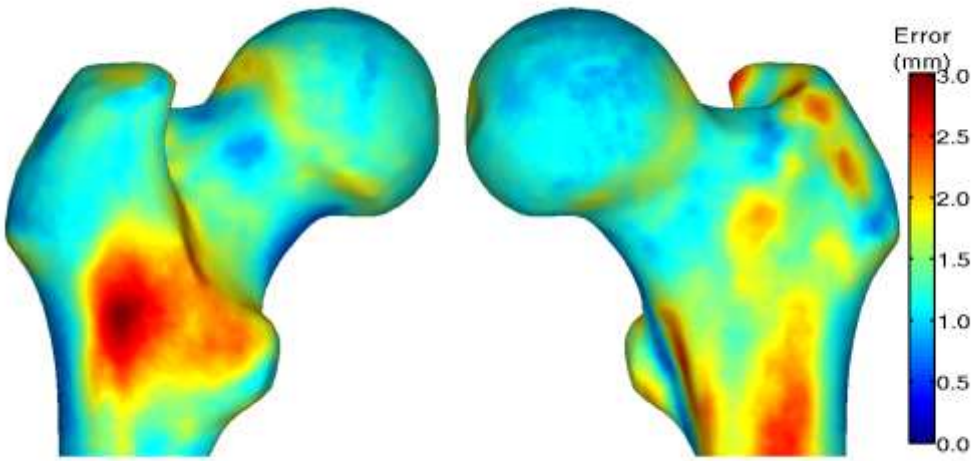


Figure 5: Mean surface difference between the original CT images and reconstructed shape evaluated with 12 patients in set 3. Reconstructions were made using clinical DXA images from the Prodigy scanner.

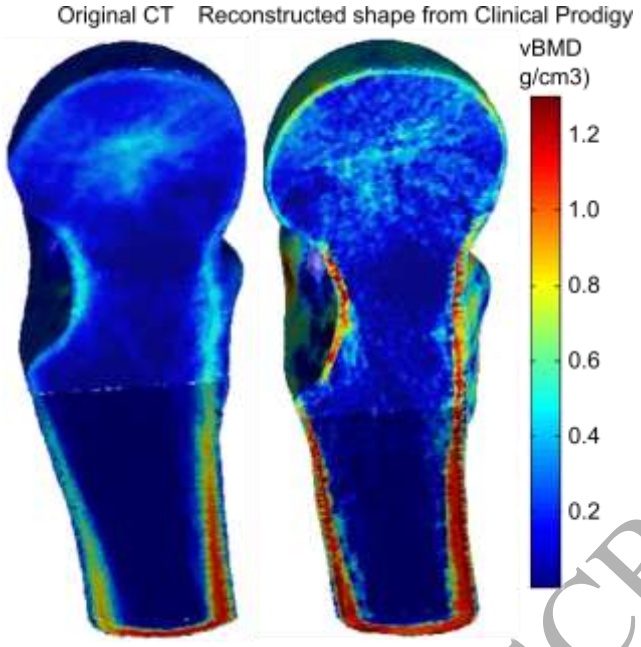


Figure 6: Representative example of the distribution of volumetric BMD from the original CT images (set 3) and the reconstructed shape using clinical DXA images from the Prodigy scanner.

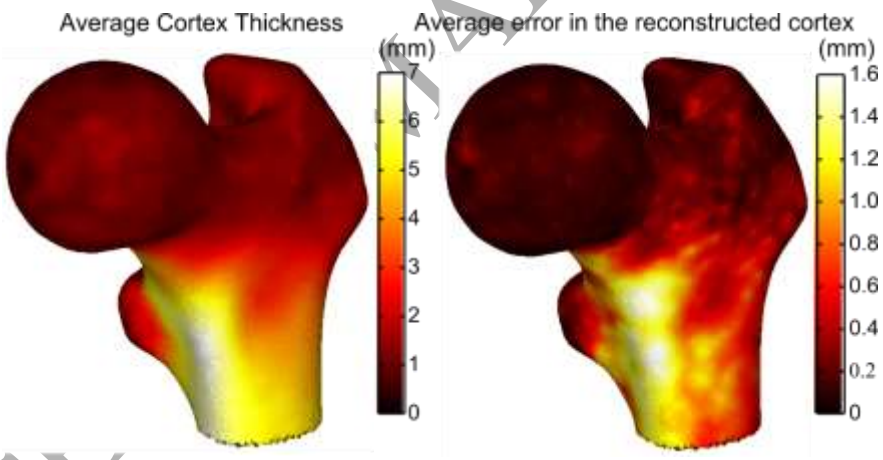


Figure 7: Average cortex thickness of the femur from original CT images and average reconstruction error of the cortex thickness evaluated with 12 patients in set 3. Reconstructions were made using clinical DXA images from the Prodigy scanner.

The linear correlation between the stiffnesses from the warped FE models and the projected CT-based FE models was r^2=0.81. The correlation between the stiffnesses from the warped and DXA-based FE models was even higher (r^2=0.85) (Fig. 8). BMD measured with DXA was also directly compared with the stiffness of the warped FE model when the resulting correlation was lower (r^2=0.72). The MAC correspondence

between the warped and CT projection-based FE models was 0.984 ± 0.011, and warped and DXA-based models was 0.977 ± 0.016.

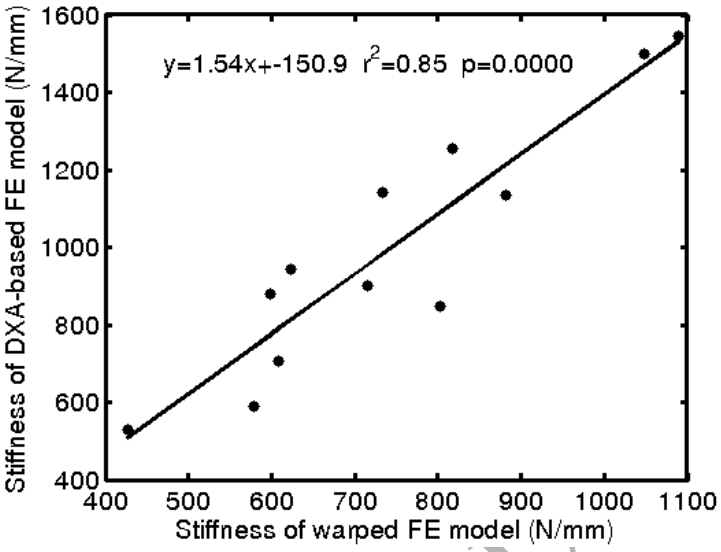


Figure 8: Stiffness between the warped and DXA-based FE models, as evaluated with CT images and DXA images of clinical subjects (set 3).

4. *Discussion*

An automatic method was developed to reconstruct the 3D shape and internal architecture of the proximal femur using only one 2D clinical DXA image. This was done by aligning a 3D statistical appearance model of the femur and pelvis with the DXA image. The outcome of the reconstruction is directly suitable for FE modelling since it produces good quality tetrahedral mesh including BMD values which can further be converted to Young's modulus. In all cases, over 99% of the elements had scaled minimum solid angle above 0.5, indicating good or excellent element quality. In addition, most of the elements with lower quality were located at the minor trochanter region, whose biomechanical relevance is minor when predicting bone strength. Those few distorted elements were automatically fixed. This means that mesh quality is not a restricting factor for the accuracy of the mechanical simulation. The presented method is also able to estimate the cortical thickness and to account for the overlaying acetabulum on the femoral head in a DXA image.

ACCEPTED MANUSCRIPT

In literature, when only one DXA or X-ray image was used for the reconstruction, the mean point-to-surface reconstruction error generally ranged between 1.1 and 3.0 mm (Galibarov et al., 2010; Hurvitz and Joskowicz, 2008; Langton et al., 2009a; Thevenot et al., 2013; Whitmarsh et al., 2011). Thus, the current findings (1.0 and 1.4 mm) are among the most accurate reported. Compared to our earlier method (Väänänen et al., 2012b) the mean point-to-surface distance during reconstruction of cadaver femurs was slightly increased (from 0.88 mm to 1.02 mm), although in general SAM-based methods (Hurvitz and Joskowicz, 2008; Whitmarsh et al., 2011, 2010) seem to perform the reconstruction more accurately than general template-based methods (Galibarov et al., 2010; Langton et al., 2009a; Thevenot et al., 2013). The advantages in our current method compared to our earlier ones (Väänänen et al., 2012b, 2011) are that: 1) it accounts for that the acetabulum covers the femur in the DXA images and thus removes the error; 2) it automatically accounts for the orientation of the femur in the 2D DXA image since the rotations of the femur are among the optimized parameters, and 3) the outcome is directly applicable for FE simulations.

The accuracy of the vBMD density was 185mg/cm^3 which is approximately 13% of the total variation of the vBMD. The vBMD was evaluated element-by-element between the mesh of the original CT images and reconstructed shape. Our average element volume was less than one tenth of the voxel volume used earlier in the voxel-by-voxel comparison (Whitmarsh et al., 2011). This partly explains the larger variation in vBMD compared to what has been reported earlier. Standardized evaluation protocols would help to compare the methods more consistently (Markelj et al., 2012).

Table 2 and Fig. 4 present how the errors accumulate when the reconstruction procedure advances from simplified ones towards the more complex clinical case. The largest single source of the error (0.47 mm, i.e., 33 % of the error in the clinical case) in the reconstruction of the femoral shape (i.e., in the point-to-surface difference) seems to be the creation of the femoral SAM. In addition, the 2D-to-3D conversion accumulated 0.40 mm to the error i.e., 27% of the error in the clinical case. By contrast, using noisier DXA image instead of an optimal projected CT image seems to affect the reconstruction error non-significantly. Similarly, using newer DXA scanner (Lunar iDXA) with higher resolution resulted in similar accuracy as the older DXA device (Lunar Prodigy). Thus, all reference images contained enough information about the 2D shape for reconstruction and other sources of error were dominant, such as lack of information in the third direction.

ACCEPTED MANUSCRIPT

Non-ideal training set (Recreation vs. Leave-one-out) accumulated error of approximately 0.16mm (12 % of the error in the clinical case). Using images of clinical subjects, instead of cadaver femurs, as the reference image accumulated 0.35 mm to the error (25% of the error in the clinical case).

In the DXA images from clinical patients, potential additional sources of errors are lower signal-to-noise ratio due to soft tissue coverage, and the pelvis which covers the femoral head. In our study, also the demographics parameters (age and sex) differed between the training sets and the test set. In addition, the resolution, to which the reconstructed shape is compared, was lower in the original CT images of the clinical subjects (restriction in our ethical permission) than in the CT images of the training set. The number of subjects in both training and evaluations sets were relatively low (34 and 12). However, the method shows promising performance already with these limited sets. Therefore, to improve the training of SAMs, it is necessary in the future to collect more femurs from patients with similar ethnical parameters as the patients in our clinical test set. Now it remains unclear if the difference in the reconstruction accuracy between our method and the most accurate one reported earlier (Whitmarsh et al., 2011) was due to the gender differences between the test and train sets, the lower resolution in the CT images of the test set, or due to differences between the methods. Adding SAM of the pelvis to the reconstruction of the shape positively influenced the accuracy in the femoral head, as compared to that of earlier studies (Whitmarsh et al., 2011). The head and neck were the most accurately reconstructed regions in our study. The head was estimated with 40% higher accuracy than that reported by Whitmarsh et al. (2011). These authors accounted for the overlaying pelvis in DXA images by extrapolating the femoral head region based on the other regions. Adding the pelvis to the reconstruction has also another benefit since it allows more realistic boundary conditions in the FE simulations in the future.

Correlation of the stiffness between the QCT-based and DXA-based models was r^2=0.85. This is in agreement or better than earlier findings (Langton et al., 2009b; Väänänen et al., 2012b) and shows good agreement between the models. The MAC value between the models was 0.977, which manifests good agreement between the predicted local displacements. Therefore, the present results are encouraging and suggest close agreement between the FE models of the reconstructed shapes and the results from experimental tests in the future (Grassi et al., 2013; Väänänen et al., 2013).

ACCEPTED MANUSCRIPT

Modelling of the cortical bone separately in the FE model, based on direct segmentation from clinical CT images, has earlier been shown to improve the prediction of the experimental femoral strength (Koivumäki et al., 2012a, 2012b). Our study demonstrates that the 3D cortical thickness can be reconstructed from one DXA image with good accuracy. However, since a threshold value was given for the minimum cortical thickness, current methodology may slightly overestimate the reconstruction accuracy of the cortical thickness at regions with very thin cortex such as in the region of the femoral head. The overestimation could be resolved in the future by utilizing cortex-thickness mapping and estimation algorithms (Treece et al., 2010, 2012).

A few manual landmarks were needed on the reference image. Placing these landmarks takes in all cases less than half a minute for an image, and the method is robust against small differences in the position of these landmarks. However, our future goal is still to make the manual positioning of these landmarks unnecessary. Calculation time for a clinical DXA-based reconstruction was on one core of Intel Sandy Bridge 2,6 GHz processor around 40 hours. Solving one FE simulation on Intel Core i7-3770 CPU @ 3.40 GHz took around 20 minutes. In the future, computation time can be substantially reduced by e.g. using graphical processor unit calculations in the generation of DRR (Ehlke et al., 2013).

In conclusion, the presented method was able to reconstruct the femoral shape, cortical thickness and internal density based on one clinical DXA image with high accuracy (e.g. shape error of 1.4mm). It was also able to automatically generate 3D meshes which are suitable for FE analysis. DXA-based FE analyses agreed well with CT-based FE analyses (i.e., the warped models). Step-by-step evaluation of the reconstruction of the 3D shape showed that the generation of the SAM, reconstruction based on 2D image and utilizing clinical DXA images where the most significant sources of error in the reconstruction, whereas use of noisier DXA images instead of projection of CT images as reference image seemed to have only minor effects.

Acknowledgement

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

The authors acknowledge CSC- the Finnish IT Center for Science for computational tools, and the Swedish Research Council (2011-5064), the Crafoord foundation, the North-Savo cultural foundation and National Doctoral Programme of Musculoskeletal Disorders and Biomaterials for funding.

Appendix

Supplementary material contains more detailed description of the landmarks on the femur and pelvis. Please see supplementary figures S1, S2, and S3.

Vitae

Sami P. Väänänen is currently a visiting post-doc at Lund University, Sweden. He received his PhD degree in medical physics at University of Eastern Finland, Kuopio, Finland in 2014. He received his MSc degree in computational physics from the same university in 2009. He focuses on functional imaging of bone for diagnosis of osteoporosis by utilizing techniques in image and statistical shape analysis, mechanical simulation and mechanical testing.

Lorenzo Grassi is currently a PhD student at the Department of Biomedical Engineering, Lund University, Sweden. He received his M.Sc. in Biomedical Engineering at University of Pisa, Italy, in 2008. After that, he joined the Medical Technology Lab of the Rizzoli Orthopaedic Institute in Bologna, Italy, which he left in 2011 to start his current position. His research interests include finite element modelling of human bones aimed at estimation of bone strength and fracture risk. He has been also active in the field of mesh morphing, and statistical modelling of bone shape and mineral density.

Gunnar Flivik is associate professor and senior consultant orthopaedic surgeon at Skåne University Hospital, Lund University, Sweden. He received his PhD in 2005 in Orthopaedics from Lund University and was then working as a researcher and orthopaedic surgeon in Perth, Australia. Since 2007 he is back in Lund involved mainly in joint arthroplasty both clinically and scientifically. He is heading the Radio Stereometry Analysis (RSA) laboratory in Lund.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Jukka S. Jurvelin is a professor in medical physics at the Department of Applied Physics, University of Eastern Finland. He received his PhD degree in 1993 from University of Kuopio, Finland, and worked as a postdoctoral researcher from 1993 to 1995 at Mueller Institute for Biomechanics, Bern, Switzerland. His research interests include development of quantitative biomechanical and imaging methods for sensitive diagnostics of osteoporosis and osteoarthritis.

Hanna Isaksson is an associate professor in biomechanics at the Department of Biomedical Engineering, Lund University, Sweden. She received her PhD degree in biomedical engineering from Eindhoven University of Technology in 2007, and was thereafter a postdoctoral researcher at University of Eastern Finland, Kuopio, Finland between 2008 to 2011. Her research focuses on the assessment of bone strength and bone quality using imaging, spectroscopic, and diffraction methods in combination with computational modelling.

ACCEPTED MANUSCRIPT

References

Ahmad, O.M., Ramamurthi, K., Wilson, K.E., Engelke, K., Prince, R.L., Taylor, R.H., 2010. Volumetric DXA (VXA): A new method to extract 3D information from multiple in vivo DXA images. J. bone Miner. Res. 25, 2744–2751.

Allemang, R.J., 2003. The modal assurance criterion-twenty years of use and abuse. Sound Vib. 37, 14–23.

Allemang, R.J., Brown, D., 1982. A correlation coefficient for modal vector analysis, in: Proceedings, International Modal Analysis Conference. pp. 110–116.

André, B., Dansereau, J., Labelle, H., 1994. Optimized vertical stereo base radiographic setup for the clinical three-dimensional reconstruction of the human spine. J. Biomech. 27, 1023–35.

ANSYS, 2012. Mechanical APDL Theory Reference Release 14.5. Southpointe 275 Technology Drive Canonsburg, PA 15317.

Bade, R., Haase, J., Preim, B., 2009. Comparison of fundamental mesh smoothing algorithms for medical surface models, SimVis. 6, 289-304.

Bonnick, S. Lou, 2007. HSA: beyond BMD with DXA. Bone 41, S9–S12. doi:10.1016/j.bone.2007.03.007

Bookstein, F.L., 1989. Principal warps: Thin-plate splines and the decomposition of deformations. Pattern Anal. Mach. Intell. IEEE Trans. 11, 567–585.

ACCEPTED MANUSCRIPT

Caponetti, L., Fanelli, A.M., 1990. 3D Bone reconstruction from two X-Ray views. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 12, 208–210.

Cauley, J.A., 2013. Public health impact of osteoporosis. J. Gerontol. A. Biol. Sci. Med. Sci. 68, 1243–1251.

Chernoff, K., Nielsen, M., 2009. Projected generalized procrustes alignment. Inf. Process. Med. Imaging 21, 503–14.

Chevalier, Y., Pahr, D., Zysset, P.K., 2009. The role of cortical shell and trabecular fabric in finite element analysis of the human vertebral body. J. Biomech. Eng. 131, 111003.

Cody, D.D., Gross, G.J., Hou, F.J., Spencer, H.J., Goldstein, S.A., Fyhrie, D.P., 1999. Femoral strength is better predicted by finite element models than QCT and DXA. J. Biomech. 32, 1013–1020.

Cootes, T.F., Edwards, G.J., Taylor, C.J., 2001. Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. 23, 681–685.

Ehlke, M., Ramm, H., Lamecker, H., Hege, H.-C., Zachow, S., 2013. Fast generation of virtual X-ray images for reconstruction of 3D anatomy. IEEE Trans. Vis. Comput. Graph. 19, 2673–82.

Fang, Q., Boas, D.A., 2009. Tetrahedral mesh generation from volumetric binary and grayscale images, in: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, pp. 1142–1145.

Fleute, M., Lavallée, S., 1999. Nonrigid 3-d/2-d registration of images using statistical models, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI'99. Springer, pp. 138–147.

Ford, C.M., Keaveny, T.M., Hayes, W.C., 1996. The effect of impact direction on the structural capacity of the proximal femur during falls. J. Bone Miner. Res. 11, 377–83.

Galibarov, P.E., Prendergast, P.J., Lennon, A.B., 2010. A method to reconstruct patient-specific proximal femur surface models from planar pre-operative radiographs. Med. Eng. Phys. 32, 1180–1188.

Gower, J.C., 1975. Generalized procrustes analysis. Psychometrika 40, 33–51.

Grassi, L., Hraiech, N., Schileo, E., Ansaloni, M., Rochette, M., Viceconti, M., 2011. Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur. Med. Eng. Phys. 33, 112–20.

Grassi, L., Schileo, E., Taddei, F., Zani, L., Juszczyk, M., Cristofolini, L., Viceconti, M., 2012. Accuracy of finite element predictions in sideways load configurations for the proximal human femur. J. Biomech. 45, 394–399.

Grassi, L., Väänänen, S.P., Amin Yavari, S., Weinans, H., Jurvelin, J.S., Zadpoor, A.A., Isaksson, H., 2013. Experimental validation of finite element model for proximal composite femur using optical measurements. J. Mech. Behav. Biomed. Mater. 21, 86–94.

Holzer, G., Von Skrbensky, G., Holzer, L.A., Pichl, W., 2009. Hip fractures and the contribution of cortical versus trabecular bone to femoral neck strength. J. bone Miner. Res. 24, 468–474. doi:10.1359/JBMR.081108

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Humbert, L., Carlioz, H., Baudoin, A., Skalli, W., Mitton, D., 2008. 3D Evaluation of the acetabular coverage assessed by biplanar X-rays or single anteroposterior X-ray compared with CT-scan. Comput. Methods Biomech. Biomed. Engin. 11, 257–62.

Humbert, L., Whitmarsh, T., Craene, M. De, Del Río Barquero, L.M., Frangi, A.F., 2012. Technical note: comparison between single and multiview simulated DXA configurations for reconstructing the 3D shape and bone mineral density distribution of the proximal femur. Med. Phys. 39, 5272–5276.

Humbert, L., Whitmarsh, T., De Craene, M., del Rio Barquero, L.M., Fritscher, K.D., Schubert, R., Eckstein, F., Link, T., Frangi, A.F., 2010. 3D reconstruction of both shape and Bone Mineral Density distribution of the femur from DXA images, in: 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Ieee, pp. 456–459.

Hurvitz, A., Joskowicz, L., 2008. Registration of a CT-like atlas to fluoroscopic X-ray images using intensity correspondences. Int. J. Comput. Assist. Radiol. Surg. 3, 493–504.

Iwashita, Y., Kurazume, R., Nakamura, K., Okada, T., Sato, Y., 2007. In-vivo experiments of 3D femoral shape estimation using two 2D fluoroscopic images, in: Robotics. pp. 3008–3008.

Kanis, J.A., 2002. Diagnosis of osteoporosis and assessment of fracture risk. Lancet 359, 1929–1936.

Kanis, J.A., Johnell, O., 2005. Requirements for DXA for the management of osteoporosis in Europe. Osteoporos. Int. 16, 229–238.

Kanis, J.A., Melton, L.J., Christiansen, C., Johnston, C.C., Khaltaev, N., 1994. The diagnosis of osteoporosis. J. Bone Miner. Res. 9, 1137–41.

Karjalainen, J.P., Riekkinen, O., Töyräs, J., Hakulinen, M., Kröger, H., Rikkonen, T., Salovaara, K., Jurvelin, J.S., 2012. Multi-site bone ultrasound measurements in elderly women with and without previous hip fractures. Osteoporos. Int. 23, 1287–1295.

Knaan, D., Joskowicz, L., 2003. Effective Intensity-Based 2D / 3D Rigid Registration between Fluoroscopic X-Ray and CT, in: Medical Image Computing and Computer-Assisted Intervention - MICCAI International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 351–358.

Koivumäki, J.E.M., Thevenot, J., Pulkkinen, P., Kuhn, V., Link, T.M., Eckstein, F., Jämsä, T., 2012a. Cortical Bone Finite Element Models in the Estimation of Experimentally Measured Failure Loads in the Proximal Femur. Bone 51, 737–740.

Koivumäki, J.E.M., Thevenot, J., Pulkkinen, P., Kuhn, V., Link, T.M., Eckstein, F., Jämsä, T., 2012b. Ct-based finite element models can be used to estimate experimentally measured failure loads in the proximal femur. Bone 50, 824–9.

Kolta, S., Le Bras, A., Mitton, D., Bousson, V., de Guise, J. a, Fechtenbaum, J., Laredo, J.D., Roux, C., Skalli, W., 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.

Kurazume, R., Nakamura, K., Okada, T., Sato, Y., Sugano, N., Koyama, T., Iwashita, Y., Hasegawa, T., 2009. 3D reconstruction of a femoral shape using a parametric model and two 2D fluoroscopic images. Comput. Vis. Image Underst. 113, 202–211.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Langton, C.M., Pisharody, S., Keyak, J.H., 2009a. Generation of a 3D proximal femur shape from a single projection 2D radiographic image. Osteoporos. Int. 20, 455–61.

Langton, C.M., Pisharody, S., Keyak, J.H., 2009b. 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.

Larsson, D., Luisier, B., Kersh, M.E., Dall'ara, E., Zysset, P.K., Pandy, M.G., Pahr, D.H., 2014. Assessment of transverse isotropy in clinical-level CT images of trabecular bone using the gradient structure tensor. Ann. Biomed. Eng. 42, 950–9.

Link, T.M., 2012. Osteoporosis imaging: state of the art and advanced imaging. Radiology 263, 3–17.

Liu, A., Joe, B., 1994. Relationship between tetrahedron shape measures. BIT Numer. Math. 34, 268–287.

López, E., Ibarz, E., Herrera, A., Mateo, J., Lobo-Escolar, A., Puértolas, S., Gracia, L., 2012. A mechanical model for predicting the probability of osteoporotic hip fractures based in DXA measurements and finite element simulation. Biomed. Eng. Online 11, 84.

Luo, Y., Ferdous, Z., Leslie, W., 2013. Precision study of DXA based patient-specific finite element modeling for assessing hip fracture risk. Int. j. numer. method. biomed. eng. 29, 615–629.

Markelj, P., Tomaževič, D., Likar, B., Pernuš, F., 2012. A review of 3D/2D registration methods for image-guided interventions. Med. Image Anal. 16, 642–661.

Melton, L.J., Chrischilles, E.A., Cooper, C., Lane, A.W., Riggs, B.L., 1992. How Many Women Have Osteoporosis? J. Bone Miner. Res. 7, 1005–1010.

Messmer, P., Long, G., Suhm, N., Regazzoni, P., Jacob, a L., 2001. Volumetric model determination of the tibia based on 2D radiographs using a 2D/3D database. Comput. Aided Surg. 6, 183–94.

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

Naylor, K.E., McCloskey, E. V, Eastell, R., Yang, L., 2013. Use of DXA-based finite element analysis of the proximal femur in a longitudinal study of hip fracture. J. Bone Miner. Res. 28, 1014–21.

Pasco, J.A., Seeman, E., Henry, M.J., Merriman, E.N., Nicholson, G.C., Kotowicz, M.A., 2006. The population burden of fractures originates in women with osteopenia, not osteoporosis. Osteoporos. Int. 17, 1404–1409.

Sarkalkan, N., Waarsing, J.H., Bos, P.K., Weinans, H., Zadpoor, A.A., 2014. Statistical shape and appearance models for fast and automated estimation of proximal femur fracture load using 2D finite element models. J. Biomech. 47, 3107–14.

Schileo, E., Balistreri, L., Grassi, L., Cristofolini, L., Taddei, F., 2014. To what extent can linear finite element models of human femora predict failure under stance and fall loading configurations? J. Biomech. 47, 3531–8.

Taddei, F., Schileo, E., Helgason, B., Cristofolini, L., Viceconti, M., 2007. The material mapping strategy influences the accuracy of CT-based finite element models of bones: an evaluation against experimental measurements. Med. Eng. Phys. 29, 973–979.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Thevenot, J., Koivumäki, J.E.M., Kuhn, V., Eckstein, F., Jämsä, T., 2013. A novel methodology for generating 3D finite element models of the hip from 2D radiographs. J. Biomech. 47, 438–444.

Trabelsi, N., Yosibash, Z., 2011. Patient-specific finite-element analyses of the proximal femur with orthotropic material properties validated by experiments. J. Biomech. Eng. 133, 061001.

Treece, G.M., Gee, a. H., Mayhew, P.M., Poole, K.E.S., 2010. High resolution cortical bone thickness measurement from clinical CT data. Med. Image Anal. 14, 276–290.

Treece, G.M., Poole, K.E.S., Gee, a H., 2012. Imaging the femoral cortex: thickness, density and mass from clinical CT. Med. Image Anal. 16, 952–65.

Whitmarsh, T., Fritscher, K., 2011. A statistical model of shape and bone mineral density distribution of the proximal femur for fracture risk assessment. Med. Image Comput. Comput. Interv. – MICCAI 2011 Lect. Notes Comput. Sci. 6892, 393–400.

Whitmarsh, T., Fritscher, K.D., Humbert, L., del Río Barquero, L.M., Roth, T., Kammerlander, C., Blauth, M., Schubert, R., Frangi, A.F., 2012. Hip fracture discrimination from dual-energy X-ray absorptiometry by statistical model registration. Bone 51, 896–901.

Whitmarsh, T., Humbert, L., De Craene, M., Del Rio Barquero, L.M., Frangi, A.F., 2011. Reconstructing the 3D Shape and Bone Mineral Density Distribution of the Proximal Femur from Dual-energy X-ray Absorptiometry. IEEE Trans. Med. Imaging 30, 2101–2114.

Whitmarsh, T., Humbert, L., De Craene, M., del Río Barquero, L.M., Fritscher, K., Schubert, R., Eckstein, F., Link, T., Frangi, A.F., 2010. 3D Bone Mineral Density Distribution and Shape Reconstruction of the Proximal Femur from a Single Simulated DXA Image: An In Vitro Study, in: Dawant, B.M., Haynor, D.R. (Eds.), Proceedings of SPIE Medical Imaging 2010: Image Processing. p. 76234U.

Whitmarsh, T., Humbert, L., del Rio Barquero, L.M., Di Gregorio, S., Frangi, A.F., Del Río Barquero, L.M., 2013. 3D reconstruction of the lumbar vertebrae from anteroposterior and lateral dual-energy X-ray absorptiometry. Med. Image Anal. 17, 475–487.

WHO, 1994. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. WHO Tech. Rep. Ser. 843, 1–129.

Viceconti, M., Taddei, F., Montanari, L., Testi, D., Leardini, A., Clapworthy, G., Van Sint Jan, S., 2007. Multimod Data Manager: a tool for data fusion. Comput. Methods Programs Biomed. 87, 148–159.

Väänänen, S.P., Amin Yavari, S., Weinans, H., Zadpoor, A.A., Jurvelin, J.S., Isaksson, H., 2013. Repeatability of digital image correlation for measurement of surface strains in composite long bones. J. Biomech. 46, 1928–32. doi:10.1016/j.jbiomech.2013.05.021

Väänänen, S.P., Isaksson, H., Julkunen, P., Sirola, J., Kröger, H., Jurvelin, J.S., 2011. Assessment of the 3-D shape and mechanics of the proximal femur using a shape template and a bone mineral density image. Biomech. Model. Mechanobiol. 10, 529–38.

Väänänen, S.P., Isaksson, H., Waarsing, E., Zadpoor, A.A., Jurvelin, J.S., Weinans, H., 2012a. Estimation of 3D rotation of femur in 2D hip radiographs. J. Biomech. 45, 2279–2283.

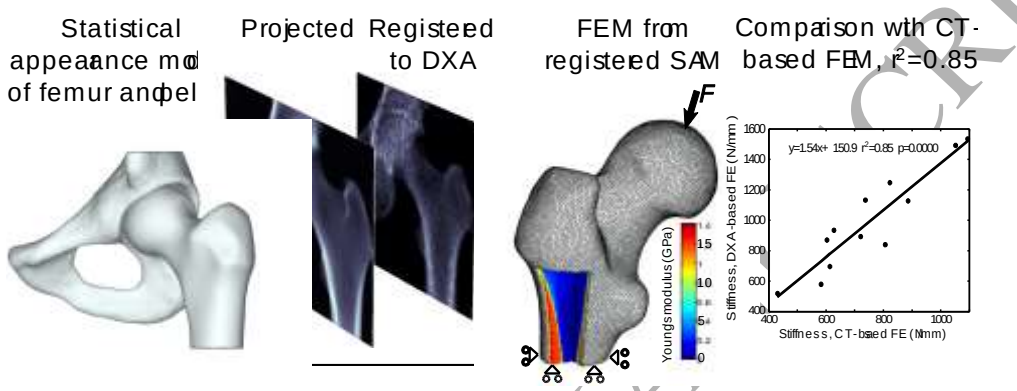
Väänänen, S.P., Jurvelin, J.S., Isaksson, H., 2012b. Estimation of 3-D shape, internal density and mechanics of proximal femur by combining bone mineral density images with shape and density templates. Biomech. Model. Mechanobiol. 11, 791–800

ACCEPTED MANUSCRIPT

Väänänen, S.P., Jurvelin, J.S., Isaksson, H., 2012c. Estimation of 3D shape and internal density of proximal femur by combining DXA images with shape and density templates, in: Transactions of the Orthopaedic Research Society, San Francisco.

Zheng, G., 2009. Statistical deformable model-based reconstruction of a patient-specific surface model from single standard x-ray radiograph, in: Computer Analysis of Images and Patterns. Springer, pp. 672–679.

Graphical_abstract



ACCEPTED MANUSCRIPT