Fast Prediction of Femoral Biomechanics Using Supervised Machine Learning and Statistical Shape Modeling

Elham Taghizadeh, Michael Kistler, Philippe Büchler and Mauricio Reyes

University of Bern, Institute for Surgical Technology and Biomechanics, Switzerland

{elham.taghizadeh; michael.kistler; philippe.buechler; mauricio.reyes}@istb.unibe.ch

http://www.istb.unibe.ch/

Abstract Finite element (FE) analysis is an important computational tool in biomechanics. However, its adoption into clinical practice has been hampered by its computational complexity and required high technical competences for clinicians. In this paper we propose a supervised learning approach to predict the outcome of the FE analysis. We demonstrate our approach on clinical CT and X-ray femur images for FE predictions (*FEP*), with features extracted, respectively, from a statistical shape model and from 2D-based morphometric and density information. Using leave-one-out experiments and sensitivity analysis, comprising a database of 89 clinical cases, our method is capable of predicting the distribution of stress values for a walking loading condition with an average correlation coefficient of 0.984 and 0.976, for CT and X-ray images, respectively. These findings suggest that supervised learning approaches have the potential to leverage the clinical integration of mechanical simulations for the treatment of musculoskeletal conditions.

Keywords Finite Element Analysis, Biomechanics, Statistical Shape Modeling, Machine Learning, Femur

Introduction

Osteoporosis is a very frequent disease that affects the life of many people after the age of 50. Osteoporosis causes annually more than 2.3 million fractures in Europe and in the USA. In 2002, it was reported that in England and Wales, the osteoporosis related fractures cost £942 million annually and this value would increase with the ageing of the population in the western countries [1]. An accurate estimation of bone strength and fracture risk can help the diagnosis of osteoporosis, leading to an improvement of patient's quality of life and reduced associated healthcare costs.

Dual energy X-ray absorptiometry (DEXA) scan is the standard clinical diagnostics tool to evaluate the level of osteoporosis and the related risk of fracture. A T-score smaller than or equal to -2.5 of femoral neck or lumbar spine indicates osteoporosis. T-score is the number of standard deviations (STD) that bone mineral density deviates from the average of bone mineral density (BMD), measured in a healthy 30 year old population with the same gender and ethnicity as the patient [1].

To automate the diagnosis of osteoporosis from DEXA images, Whitmarsh and colleagues used statistical shape and appearance models. They proposed a Fisher Linear Discriminant Analysis (FLDA) method to classify bones having a high or low fracture risk [2]. Sarkalkan and colleagues proposed 2D finite element models built from DEXA images to predict the fracture risk of the proximal femur [3].

It has been shown that 3D (FE) analyses predict bone strength more accurately than clinical methods such as DEXA [4]. However, the adoption of FE analyses into clinical practice has been hampered by its computational complexity and required technical competences. To analyze the bone behavior under a certain loading condition, an accurate segmentation of the bone is necessary, a valid finite element mesh must be generated, and appropriate boundary conditions need to be applied to the model. These preparation steps are followed by time-consuming calculations to determine the biomechanical behavior of the bone. All of these steps are time consuming and computationally demanding, which make the FE analysis less appealing for clinicians [5]. As a consequence, up to now, the FE analysis techniques did not reach the clinical workflow. Different research studies aimed to automate the segmentation [6] and finite element mesh creation [7,8], however to the best of our knowledge no method has been proposed to bypass the computational complexity of FE calculations. In this paper we aim at alleviating the aforementioned issues of FE analysis to promote their adoption into clinical practice.

The two most important features describing bone biomechanics are shape and bone mineral density (BMD). Therefore, we hypothesize that machine-learning techniques can be used to predict the biomechanical properties of the bone using shape and density features extracted from clinical patient scans, as well as patient anthropometric information. To this end, we propose a supervised learning approach to predict the outcome of FE analysis. As feature predictors for bone shape and density, we propose to characterize this information in a compact way by using statistical shape modeling of the anatomy [6]. In this way, we reduce the dimensionality of the feature space, which leverages the building process of the machine learning model, and moreover, allows us to exploit previous developments in statistical shape modeling (e.g. active shape models [9]). We demonstrate this by predicting bone stresses from clinical CT (*FEP*), where features are extracted from a statistical shape and statistical intensity model of the human femur and patient's anthropometric information. As a second demonstration we present preliminary results on a simplified scenario where FE femur biomechanics are predicted from 2D X-ray images. Morphometric and density information available in the 2D image was used as predictors.

In the next section we describe in detail how the prediction models are built, how features are defined and extracted, and one example scenario to demonstrate how the approach can be adapted for X-ray scans. In the Results section, the databases used for training and testing of the method are presented and the quality of the prediction is quantified. We conclude the paper by discussing the advantages and limitations of our proposed approach.

Method

In this section we explain the proposed method for finite element prediction, termed here *FEP*. We then follow by exemplifying how the proposed method can be employed for a different image modality, such as X-ray.

Finite Element Prediction (FEP) Framework

The main framework for *FEP* is summarized in Figure 1. Following the same scheme as in supervised learning, our approach has two stages.

First, during the training stage, a statistical model of shape and intensity is created as in [10]. In short, an iterative mesh morphing method [11] is used to compute point correspondences for a dense volumetric mesh consisting of approximately 190,000 nodes and 130,000 tetrahedral elements. Bone density for each node is then extracted from the original CT scans [10]. A principal component analysis (PCA) is then performed separately on shape and density information, yielding two separate models. As shown in Figure 1 (training phase) each bone can then be modeled through shape and density scores.

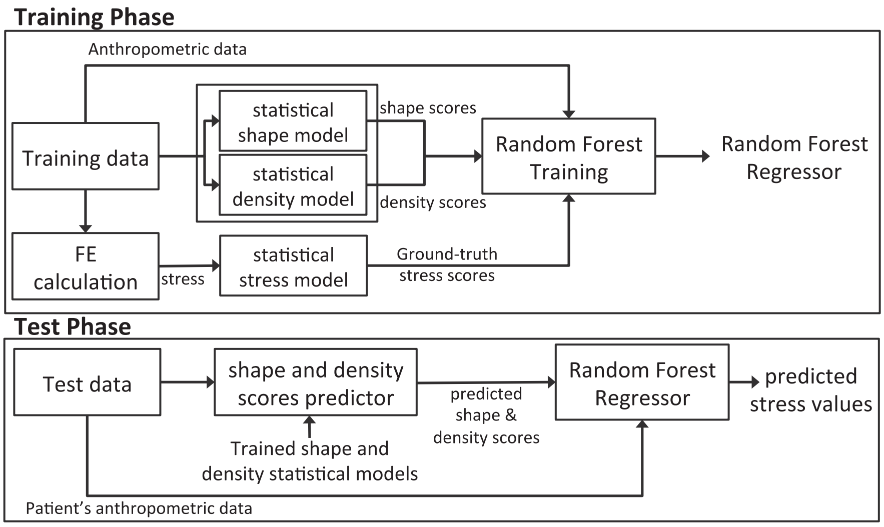

Figure 1: Schematic description of the *FEP* predictions. Using shape, density, and stress scores of training data, we learn a random forest regression model. In the test phase, the trained random forest predicts from anthropometric and SSM-based bone and density predictors, extracted for a new image, the parameters of the statistical stress model.

As response variables, FE computations are used to calculate stress values on each node of the FE mesh. The FE analyses were performed with the commercial package Abaqus/Standard (Abaqus v6.12, SIMULIA, USA). Boundary conditions (BC) representing a walking situation were applied to the bone models. We chose the loadings of the joint configuration proposed in [12], where the node constraints are selected at the femoral head, the intercondylar femoral notch, and the lateral epicondyle of the femur. The force values were calculated based on the body weight. The calculated nodal stress values were used to build a statistical model of stress.

A statistical model of the stress in the model was built. The scores of this model were used as output of the prediction algorithm. For the calculations, we considered only the top modes of shape, density and stress obtained from the statistical models. The number of modes included was based on the criterion to keep 98% of the variation that was in the dataset.

Using the set of aforementioned predictors and response variables, a random-forest model [13] was trained to work as the regressor. Random forests are being used for different classification and regression problems [14]. They are robust to noise and more importantly are able to predict the output even when some input information is missing. Besides, Random forests are naturally conceived to use nature of different data, as here anthropometric, morphometric and BMD Information is used.

We note here that as the output of the prediction is the parameters of orthogonal vectors, it is possible to train one random forest regressor for each stress parameter. As suggested in [13], one-third of features are selected for each node-split, and the maximum depth for the tree is selected based on a 10-fold cross validation.

During the test phase, given a patient CT image of the anatomy, the feature extraction process consists of projecting the patient's anatomy into the shape space to recover shape and density parameters [6,10]. Here is where current and advanced SSM-based technologies (e.g. active shape models, hierarchical shape models [6]) can be used to compute scores for shape and density information. For the sake of simplicity we relied our experiments on a leave-one-out scheme where these parameters are extracted during model building. We also included anthropometric features such as patient's age, gender, height, and weight in the input features. Finally, after feature extraction, FE predictions can be computed to yield stress scores, which are converted into stress values by simply drawing the corresponding sample values from the statistical model of stresses.

FEP for X-ray Images

By employing statistical shape and density scores to represent the anatomy and predict bone biomechanics, it is possible to decouple the prediction model from the input image modality. In other words, bone shape and density scores act as a "bridge" connecting the image modalities used to capture bone shape and density information of the patient to the image modality (CT scan) used to characterize bone biomechanics. As an example of using *FEP* for a different image modality, we demonstrate in this paper the case of having X-ray images as the input image modality used to capture bone shape and density information. We then show how to connect this information to shape and density scores used by *FEP* to predict bone biomechanics.

For the sake of simplicity, in this study we built synthetic X-ray scans by projecting the captured CT scans to two orthogonal planes. To characterize bone shape and density information, we used a set of simple yet effective feature descriptors. From two orthogonal X-ray images a total of 21 bone morphometric dimensions, as shown in Figure 2, are extracted by selecting a few landmarks from both views. To model bone density information, the histogram of pixel intensities is calculated for the frontal view, generating a feature vector of size equal to the number of histogram bins.

From the triplets of 1) X-ray derived features, 2) patient's anthropometric data and 3) corresponding bone shape and density scores, a random forest regression model is built. During testing, a new set of previously unseen X-ray orthogonal images is used to extract morphometric and bone density features to predict the bone shape and density scores. By cascading this model with the stress prediction model, described in the previous section, we are able to perform bone biomechanics *FEP* from X-ray images.

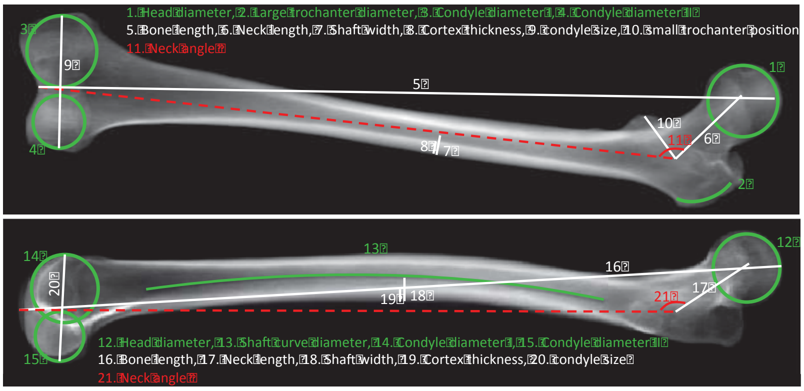

Figure 2: The morphometric feature descriptors extracted from two orthogonal views. Diameters (in green), distances (in white) and angles (in red) are shown for frontal and lateral views. By selecting three landmarks for each circle fitting and two for each line we perform the measurements.

Results

In this section we show the results of the proposed method for fast FE predictions. First the database and tools used for the study are explained. It is followed by the results of FEP method for CT and X-Ray scans.

Database and Tools

The database used in this study consists of 89 left femurs CT images. The resolution of CT scans was between 0.61 mm × 0.61 mm and 1.171 mm × 1.171 mm, with a slice thickness of 1 mm. The CT scans were acquired from 48 female and 41 male donors with average age, height, and weight of, respectively, 60.7 ± 16.2 years old, 165.70 ± 7.2 cm and 70.1 ± 13.9 kg. Table 1 reports statistics about patients and femur morphometric in our database.

Table 1: The statistics of the bones used in the database (n=89)

	Patients			Morphological parameters	
	Age	Height (cm)	Weight (kg)	Length (cm)	Anterior curve diameter (cm)
Min	23	150	42	37.8	57.0
Max	90	180	140	50.9	297.2
Mean	60.7	165.70	70.1	44.5	123.4
STD	16.2	7.2	13.9	23.2	33.4

To study the accuracy of *FEP*, we used Leave-One-Out (LOO) [15] methodology to train with the maximum number of samples. The method was tested for one sample in the database when the rest of samples were used for training. This approach was repeated until every sample in the set was tested, which resulted in 89 different sets of training and testing samples.

For each training set, we built statistical models of shape, and density [10], followed by FE computations. In the calculations, we considered the top modes of shape, density and stress statistical models with the sum of more than 98% of the variation in the dataset. As a result, 20 modes of shape and 46 modes of BMD were used for predicting the parameters of 16 modes of statistical model of stress. To tune the parameters of random forest, we performed a 10-fold cross validation using the scikit-learn toolbox [16].

Results of *FEP* for CT Scans

We evaluated the prediction accuracy of stress values for each test sample. We calculated the correlation coefficient between the ground-truth stress values for each mesh node (as generated by the FE computations, using Abaqus FE solver in the normal walking loading situation), and the predicted Mises stress for those nodes. The average correlation coefficient for 89 test cases was 0.984 with a standard deviation of 0.008, showing the high accuracy of the proposed method.

We further evaluated the prediction performance by calculating the prediction error as the difference between ground-truth and predicted stress values. The ground-truth stress values, the predicted values, and the error distribution are shown in

Figure 3 for the best and the worst results. Among 89 samples, the best result was achieved with an average error (and standard deviation) of 0.058 (0.898) MPa in the mesh. For the worst case, the average error (and standard deviation) of the predicted stress values was equal to 1.316 (7.822) MPa. After examination of the worst-case result, we found that it corresponds to a patient with a body weight of 140 Kg, while the maximum weight seen in training dataset was only 110 Kg. This can be improved by using more samples for training to cover a larger variety of the population.

To evaluate *FEP* for different regions of interest, we also examined its accuracy in the femoral neck, femoral trochanter and the femoral shaft, separately (see

Figure 4). The prediction error of stress for the neck region, which is the region of interest in fracture risk assessment, was smaller than 0.9 MPa in average.

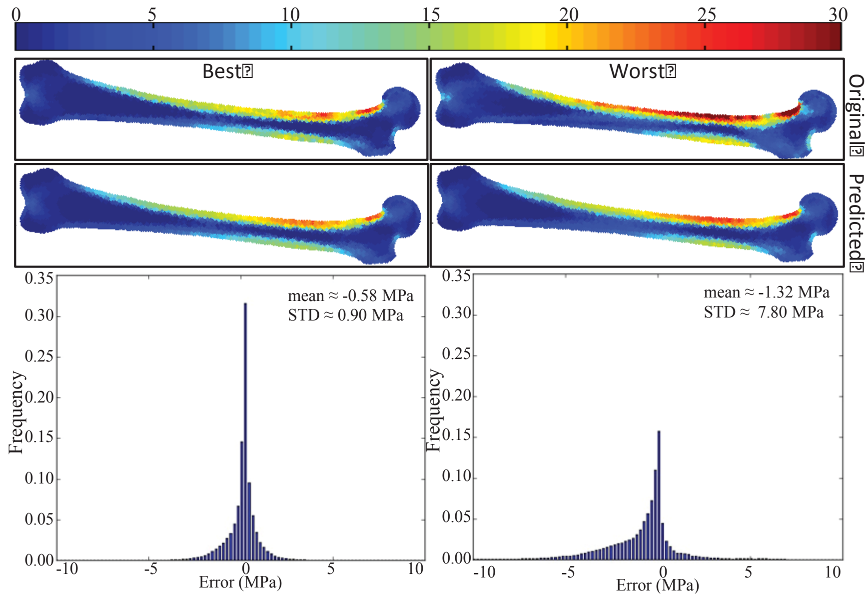

Figure 3: The stress map predicted by FEP model for the best and the worst cases. From top to bottom: the stress map calculated using FE calculations (ground-truth), the predicted stress values for the corresponding bones, the error distribution for these bones. The best result is achieved for the bone on the left column with a correlation coefficient of 0.994, and the worst prediction result in the database is in the right column with a correlation coefficient of 0.939. In absolute error distribution plots, we zoom in on the range of [-10, 10] MPa for better visibility. The frequency of error beyond this range is negligible (0.0003 and 0.0122, for the best and the worst case).

Results of *FEP* for X-ray Images

Based on features extracted from X-ray images we predicted the parameters of statistical shape and density models. We then used these parameters as input to our *FEP*. The average (standard deviation) correlation coefficient between the predicted stress using this method and the ground-truth values was 0.976 (0.012).

We developed a test case to evaluate the benefit of cascading two learning blocks (from X-ray to 3D data, and from 3D data to stress parameters) as compared to a single learning model that directly predict stresses from X-ray based features. Similarly to the other models, the depth of trees is determined based on cross-validation on training data. In this case the average correlation coefficient values dropped from 0.976 ± 0.012 to 0.956 ± 0.286. This shows that the cascading of two regression models, as proposed herein, does not significantly alter the accuracy of the prediction as compared to a single learning model. In addition, the cascading scheme has the extra value that other modality-specific models can be easily combined to *FEP*.

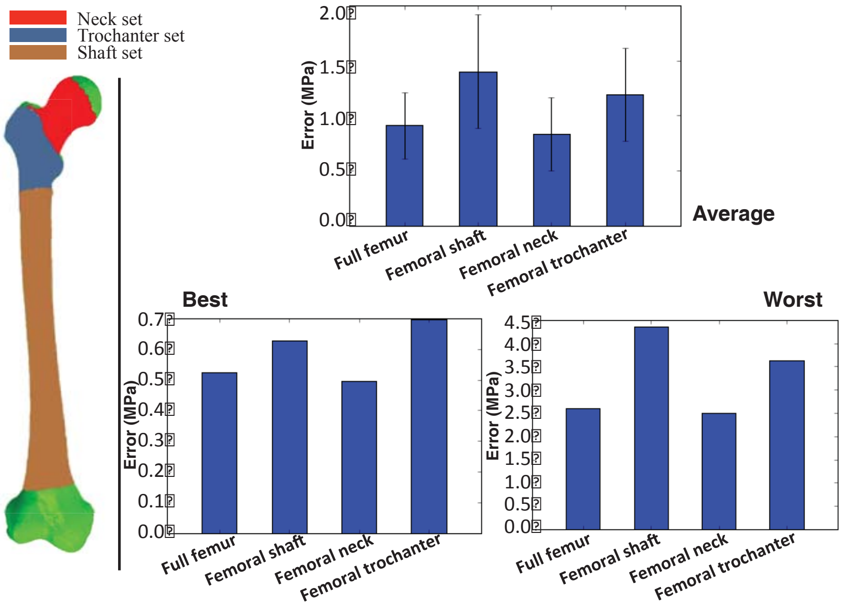

Figure 4: The absolute error of *FEP* for different parts of the bone, (top) in average, (bottom-left) best case and (bottom-right) for worst case. The different regions of interest are shown on the bone with different colors. Red: neck, blue: trochanter region and orange: shaft.

Conclusion and Discussion

It has been shown that using 3D FE analyses improves the osteoporosis diagnosis [4]. However clinical adoption of FE analysis in bone biomechanics and fracture risk assessment has been hampered by its computational complexity and required technical competences [5]. In this paper we developed a random-forest based regression framework to predict the results of the Finite Element Prediction, termed here *FEP*, by simply selecting a couple of landmarks on clinical images. We proposed to use shape and density statistical model parameters to produce a compact and predictive set of features. In addition, the approach allows other image modalities to be used for prediction, and enables the incorporation of other emerging technologies developed for statistical shape modeling.

Using leave-one-out experiments, comprising a database of 89 clinical cases, our method is capable of predicting the stress values for a walking loading condition with an average correlation coefficient of 0.984 and 0.976, for CT and X-ray images, respectively. These findings suggest that supervised learning approaches have the potential to leverage the clinical integration of mechanical simulations for the treatment of musculoskeletal conditions.

Motivated by the observed connections between the importance values obtained by random forest and actual models for shape, we analyzed the most important features in predicting the parameters of the stress statistical model. To predict the stress parameters, the body weight was found to be the most important parameter. This can be explained by the fact that in our experiments all bones were loaded in the walking situation when forces are scaled proportional to the body weight. Hence, the weight directly affects the stress values.

We note that our motivation is to demonstrate that even with rather simple, yet descriptive, selected features it is possible to yield a good level of prediction for bone biomechanics in real-time. In this work we used simple feature predictors for X-ray images. However, as proposed by the state-of-the-art approaches [17] three-dimensional bone shape and density parameters can be estimated robustly and accurately from X-ray images, which can further increase the predictive power of *FEP*.

Our method predicts the output stress values of an elastic material model for FE analysis from density and shape. However, it is flexible and can be easily adapted to incorporate more advanced mechanical parameters for predicting the bone fracture. For further improvement of *FEP*, we are planning to use existing methods on predicting trabecular bone structure from CT scans [18–20] to improve the estimation of biomechanical behavior of the bone.

This study has some limitations. First, the estimation of the scores of shape and density from CT scans was obtained using mesh registration. This registration task is time consuming and should be replaced by more effective methods such as active appearance model. However, this intermediate step is not necessary when the stress predictions are obtained from X-ray images. Another limitation results from the choice of synthetic images to mimic patients' x-ray images. This approach has been chosen to establish the method and avoid uncontrolled source of error. Clearly the accuracy of the predictions will decrease when clinical data will be used. Further studies will investigate this effect, but the high correlations reported in this study indicate that the prediction from clinical x-ray will provide accurate stress estimations. Finally, we observed that the method is not as successful in predicting the stress values for a bone of a patient who has the highest weight in our dataset. This problem occurs because no other patient with a similar body weight exists in our dataset. Similar to all other techniques that rely on machine learning, a large database that samples the population more evenly helps tackling this issue.

Our proposed approach followed by further improvements (adding trabecular bone structure to the analyses and using active shape modeling) shows a promising path towards real-time biomechanical analysis of bones in different patient-specific studies and brings an automated FE analysis to clinics. Since it is fast (the stress values are calculated in less then one second), several loading cases can be analyzed to have a better understanding of the patient's bone, moreover it can be used to find the bone strength and fracture risk for each individual patient.

The drawback of *FEP* is that for each loading case, i.e. walking, stance, side fall, one different model has to be trained. Note that this process is done offline during the training phase. The testing phase is fast and the stress values can be calculated in less than a second.

References

[1] Kanis J a. Assessment of osteoporosis at the primary health care level. World Health 2007:339.

[2] Whitmarsh T, Fritscher KD, Humbert L, Del Rio Barquero LM, Roth T, Kammerlander C, et al. A statistical model of shape and bone mineral density distribution of the proximal femur for fracture risk assessment. Med. Image Comput. Comput. Assist. Interv., vol. 14, 2011, p. 393–400.

[3] Sarkalkan N, Waarsing JH, Bos PK, Weinans H, Zadpoor AA. Statistical shape and appearance models for fast and automated estimation of

proximal femur fracture load using 2D finite element models. J Biomech 2014;47:3107–14. doi:10.1016/j.jbiomech.2014.06.027.
[4] Dall'Ara E, Luisier B, Schmidt R, Kainberger F, Zysset P, Pahr D. A nonlinear QCT-based finite element model validation study for the human femur tested in two configurations in vitro. Bone 2013;52:27–38. doi:10.1016/j.bone.2012.09.006.
[5] Poelert S, Valstar E, Weinans H, Zadpoor AA. Patient-specific finite element modeling of bones. Proc Inst Mech Eng H 2013;227:464–78. doi:10.1177/0954411912467884.
[6] Heimann T, Meinzer H-P. Statistical shape models for 3D medical image segmentation: a review. Med Image Anal 2009;13:543–63. doi:10.1016/j.media.2009.05.004.
[7] Grassi L, Schileo E, Boichon C, Viceconti M, Taddei F. Comprehensive evaluation of PCA-based finite element modelling of the human femur. Med Eng Phys 2014;36:1246–52. doi:10.1016/j.medengphy.2014.06.021.
[8] Castro-Mateos I, Pozo JM, Cootes TF, Wilkinson JM, Eastell R, Frangi AF. Statistical Shape and Appearance Models in Osteoporosis. Curr Osteoporos Rep 2014;12:163–73. doi:10.1007/s11914-014-0206-3.
[9] Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models-Their Training and Application. Comput Vis Image Underst 1995;61:38–59. doi:10.1006/cviu.1995.1004.
[10] Bonaretti S, Seiler C, Boichon C, Reyes M, Büchler P. Image-based vs. mesh-based statistical appearance models of the human femur: Implications for finite element simulations. Med Eng Phys 2014;36:1626–35. doi:10.1016/j.medengphy.2014.09.006.
[11] Boichon C, et al. shape indexing of human femur using morphing and principal component analysis. VPH, 2010.
[12] Speirs AD, Heller MO, Duda GN, Taylor WR. Physiologically based boundary conditions in finite element modelling. J Biomech 2007;40:2318–23. doi:10.1016/j.jbiomech.2006.10.038.
[13] Breiman L. Random Forests. Mach Learn n.d.;45:5–32. doi:10.1023/A:1010933404324.
[14] Criminisi A, Shotton J, editors. Decision Forests for Computer Vision and Medical Image Analysis. No. 1. Springer-Verlag London; 2013. doi:10.1007/978-1-4471-4929-3.
[15] Geisser S. Predictive Inference. CRC Press; 1993.
[16] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12:2825–30.
[17] Zheng G. 3D volumetric intensity reconsturction from 2D x-ray images using partial least squares regression. 2013 IEEE 10th Int. Symp. Biomed. Imaging, IEEE; 2013, p. 1268–71. doi:10.1109/ISBI.2013.6556762.
[18] Taghizadeh E, Maquer G, Reyes M, Büchler P. Including the trabecular anisotropy from registered microCT data in homogenized FE model improves the bone's mechanical predictions. CMBBE, Amsterdam: 2014.

[19] Larsson D, Luisier B, Kersh ME, Dall'ara E, Zysset PK, Pandy MG, et al. Assessment of transverse isotropy in clinical-level CT images of trabecular bone using the gradient structure tensor. Ann Biomed Eng 2014;42:950–9. doi:10.1007/s10439-014-0983-y.

[20] Lekadir K, Hazrati-Marangalou J, Hoogendoorn C, Taylor Z, van Rietbergen B, Frangi AF. Statistical estimation of femur micro-architecture using optimal shape and density predictors. J Biomech 2015;48:598–603. doi:10.1016/j.jbiomech.2015.01.002.