

Prediction of Biomechanical Parameters of the Proximal Femur Using Statistical Appearance Models and Support Vector Regression

Karl Fritscher¹, Benedikt Schuler¹, Thomas Link², Felix Eckstein³, Norbert Suhm⁴,
Markus Hänni⁴, Clemens Hengg⁵, and Rainer Schubert¹

¹ Institute for Biomedical Image Analysis, UMIT, Austria
{Karl.Fritscher, Benedikt.Schuler, Rainer.Schubert}@umit.at

² Department of Radiology and Biomedical Imaging
University of California, San Francisco

Thomas.Link@radiology.ucsf.edu

³ Institut für Anatomie und muskuloskelettale Forschung, Paracelsus University, Salzburg
Felix.Eckstein@pmu.ac.at

⁴ AO Development Institute, Davos, Switzerland

{Norbert.Suhm, Markus.Haenni}@aofoundation.org

⁵ Department of Trauma Surgery, Medical University Innsbruck, Austria
Clemens.Hengg@uki.at

Abstract. Fractures of the proximal femur are one of the principal causes of mortality among elderly persons. Traditional methods for the determination of femoral fracture risk use methods for measuring bone mineral density. However, BMD alone is not sufficient to predict bone failure load for an individual patient and additional parameters have to be determined for this purpose. In this work an approach that uses statistical models of appearance to identify relevant regions and parameters for the prediction of biomechanical properties of the proximal femur will be presented. By using Support Vector Regression the proposed model based approach is capable of predicting two different biomechanical parameters accurately and fully automatically in two different testing scenarios.

Keywords: Fracture load, bone strength, proximal femur, support vectors, appearance models.

1 Introduction

Fractures of the proximal femur are one of the principal causes of mortality among elderly persons. Hip fractures are primarily caused by an increased fragility of the proximal femur, due to osteoporosis or other conditions affecting bone strength. Techniques for the determination of fracture risk traditionally use methods for measuring bone mineral density (BMD) like e.g. DXA [1]. However, there is still a significant overlap in bone mineral density between osteoporotic and normal individuals and BMD alone is not sufficient to predict bone failure load for an individual patient [2].

As a consequence different other approaches were introduced using parameters based on bone geometry (e.g. [3]), trabecular (micro)-architecture (e.g. [4, 5]), and (micro)-finite element analysis (e.g. [6, 7]) to assess bone strength and estimate fracture risk.

Apart from improving the identification of patients with a high fracture risk another type of projects aims at improving the surgical treatment of femoral fractures, by predicting the risk of failure of an osteosynthesis, the so called “cut-out” risk. In [8] a tool has been developed which can be used during the surgical intervention by measuring the peak torque until complete breakaway of the cancellous bone in the femoral head. Moreover, several non invasive methods were introduced in order to predict the cut out risk using CT or x-ray images based on the calculation of Haralick texture features in CT images [9] or Minkowski functionals[10] in different regions of interest.

Beside of finding appropriate parameters most of approaches mentioned above are also dependent on defining a number of different regions of interest (ROIs) in which potentially relevant parameters are obtained. In many cases the placement of these ROIs for an individual patient needs user interaction and is therefore subjective.

Therefore in this work an approach that is based on the usage of statistical models will be presented. The key idea of this work is to use these models in order to identify relevant regions and parameters for the prediction of biomechanical properties of the proximal femur. By using these parameters as predictor variables for Support Vector Regression (=SVR) the proposed model based approach is capable of predicting two different biomechanical parameters accurately, objectively and fully automatically.

2 Methods

The workflow of the proposed algorithm can be summed up as follows:

Step 1: Create an accurate combined representation of shape and spatial intensity distribution of an object by applying an atlas based non-rigid registration on a training set of CT images (for more details refer to section 2.1)

Step 2: Apply Principal Component Analysis (=PCA) using the output of step 1 in order to reduce the dimensionality of the input space. By using a correlation-based feature selection method, a subset of n relevant Principal Components (=PCs) can be identified (for more details refer to 2.2).

Step 3: Regress the PC scores of the n relevant PCs obtained in step 2 against the biomechanical parameter of interest and find the best subset of the PCs as well as the optimal parameter values for Support Vector Regression using a stratified cross-validation scheme (for more details refer to 2.3).

Step 4: Select a subsample of the original set of input variables represented by the combined shape-intensity representations by using the eigenvector loadings of the n relevant PCs in order to reduce the number of non-relevant variables and identify relevant regions (for more details refer to 2.4).

Step 5: Repeat steps 2 and 3 with the subsample of the original set of input variables obtained in step 4.

2.1 Combined Shape-Intensity Representation

The method that is used to create a combined shape-intensity model is based on the approach presented in [11] and has been extended to generate combined representations of shape and spatial intensity-distribution [12]. The approach is based on the usage of rigid and non-rigid registration in order to align a number of subjects to an atlas subject. Performing this registration for n subjects results in n aligned intensity representations and n deformation fields $\{D_1, D_2, \dots, D_n\}$, which represent shape variations. The aligned intensity representations I_n and the shape representations D_n for each subject of the training set are placed in one vector z . The length of the vector z is $4 \times$ number of pixels within the structure of interest (e.g. the proximal femur) for 3D images.

2.2 Combined Shape-Intensity Model

Using the representations described in section 2.1, PCA is applied to reduce the dimensionality of the input data: Given n , d -dimensional training vectors, $\{z_1, \dots, z_n\}$ a training matrix M can be defined. Using Singular Value Decomposition, the covariance matrix Σ of M can be decomposed [13]. In the case when $d \gg n$, the system is underconstrained resulting in a large number of eigenvectors that will be zero. In these cases the Matrix U , containing the eigenvectors of Σ , can be calculated from the smaller Matrix T

$$T = \frac{1}{n} \sum_{i=1}^n (z_i - \mu)^T (z_i - \mu) \quad (1)$$

where μ is the mean of all training elements z [13].

The vector of PC scores α for a subject x of the training set can be calculated by using

$$\alpha = U_k^T (x - \mu) \quad (2)$$

U_k is a matrix consisting of the first k columns of U and represents the set of k orthogonal modes of shape-intensity variations. An approximation of a shape-intensity pair in 3D can then be represented by using k PCs and a k dimensional vector of PC scores α . In order to compensate the differences in scaling of the pixel intensities representing the spatial intensity distribution and the deformation field components, the pixel intensities are rescaled by a factor f in each combined shape-intensity distribution. This factor is calculated by dividing the average range between the minimum and maximum deformation of all subjects by the average range between the highest and lowest intensity value of all subject.

After performing PCA the PC scores of a subset of PCs will be used as predictor variables for SVR. This subset of PCs is determined in the next step.

2.3 Attribute Selection and Support Vector Regression

Using the PC scores and the biomechanical parameter of interest of all subjects in the training set a correlation-based feature selection method [14] with a greedy stepwise

forward search method is used in order to identify a subset of k potentially relevant PCs. N -fold cross-validation (n =number of subjects in the training set) is applied for this purpose. The PC scores of the selected k PCs will then be used to train a Support Vector Regression [15] model. In this project an exponential kernel of the form $k(x, x') = \langle x, x' \rangle^p$ with $p \in \mathbb{N}$ was used for the Support Vector Regression. For details on Support Vector Regression please refer to [15]. The best subset of the k PCs as well as the optimal values for p and for the complexity parameter C (for details on the parameters of SVR please refer to [15]) are obtained by using a stratified cross-validation scheme: First, one of the n subjects is chosen as the test data set. The remaining $n-1$ subjects serve as training set. Using this leave-one out scheme a set of parameters for the SVR parameters $C(\{0.3, 0.5, 0.7, 1, 3, 5, 10, 20, 40, 60, 80, 100\})$ and $p(\{1, 2, 3\})$ is defined. Starting with the subset of k PCs identified by the greedy stepwise forward search method, the best combination of p and C is determined by reducing the numbers of used PCs by one in each step from k PCs down to 1 PC. The order in which the PCs are excluded is determined by the ranking that results from the greedy stepwise search algorithm in the attribute selection step prior to SVR.

2.4 Reducing the Size of the Training Vectors Z_i

After the attribute selection the size of the training vectors z_i is reduced in order to remove non-relevant elements of the training vectors z_i and identify relevant properties. For this purpose the correlation coefficient r_k between the score values of the k^{th} PC for each subject in the training set and the biomechanical parameter of interest are calculated for each of the PCs that have been determined in the attribute pre-selection step. Using r_k and L_k , which is the vector containing the eigenvector loadings of the k^{th} PC, a vector U_w containing weights for each element of z_i can be calculated as follows:

$$U_w = \sum_{k=1}^m r_k |L_k| \quad (3)$$

Using U_w the dimension of each original training vector z_i is reduced by removing the elements of z_i , whose weighting values in U_w are beneath a relative threshold T . This results in n reduced training vectors zs_i . Using the vectors zs_i , a second iteration of PCA, attribute selection and support vector regression is performed to obtain a final regression model. Analogously to the determination of the optimal numbers of PCs described in section 2.3 again a stepwise procedure is used to find the best values for T in this second iteration: Initially the calculations described in sections 2.2 and 2.3 are performed setting $T=90\%$. Thereafter T is reduced by 10 % and the calculations are performed again. T is reduced stepwise by 10 % as long as the predictive quality of any of the regression models resulting from different parameter combinations for p and C does not decrease compared to the best working model in the previous step. In order to assess the predictive quality of a certain regression model the correlation between real and predicted parameter values, as well as the mean absolute error (MAE) and the relative absolute error (RAE) is measured.

2.5 Data Material

The pipeline described above was tested in two different test settings in which two biomechanical parameters should be predicted. In the first setting (=setting A) the parameter “peak torque until complete breakaway” (=PTB), which was obtained using the DensiProbe tool presented in [8], should be predicted for 14 proximal femora of men and women aged between 64 and 96 years. For this purpose CT images of cadaveric femora have been used. The values for PTB ranged from 3.9 to 12.41 Nm. The special challenge using these dataset was that the CT images were obtained using two different micro CT scanners and different image resolutions. Seven out of the 14 images had an original isotropic resolution of $0.082 \times 0.082 \times 0.082$ mm and were acquired using a Scanco DensiScan 1000 microCT scanner. The remaining seven images were acquired using a Scanco XtremeCT scanner with an original image resolution of $0.246 \times 0.246 \times 0.246$ mm. In order to simulate clinical conditions all images were resampled to images with an isotropic voxel spacing of 0.984 mm (using linear interpolation).

In test setting B, CT images of 26 femur specimen of men and women between 58 and 98 years with an original resolution of $0.2 \times 0.2 \times 0.5$ mm were used. Again – in order to simulate the resolution of images that are acquired in clinical routine - the images were resampled to images with an isotropic pixel spacing $0.7 \times 0.7 \times 0.7$ mm (using linear interpolation). The fracture load of the femora was obtained by simulating a fall on the greater trochanter (=side-impact configuration), which was leading to different types of fractures (cervical, intertrochanteric and shaft fractures) [16]. The values for the fracture load ranged from 2453 to 6359 N.

3 Results

The accuracy of the registration during the modeling process was evaluated by calculating the mean absolute error between the pixel intensities of the original CT images and the model instances that were created by using the appearance model with all PCs. Moreover, the original images and the created instance images were converted to 8-bit images (256 grey levels).

For test setting A, a statistical model of the femoral head and neck region was created. Analyzing the registration accuracy as described above a mean absolute error of 7.3 grey levels ($\sigma=6.6$), which corresponds to 2.8 % of the whole intensity range was measured. The results of the leave-one-out tests for testing scenario A are summed up in table 1. The table is showing the number of used PCs, the obtained correlation, the errors RAE and MAE, the used combination of parameters (kernel exponent and C) and the percentage of input variables that lead to the best results for 1 to k (in this case $k=7$) predictor variables (=PCs). Best results with $R=0.975$ ($p<0.01$) could be achieved using 100 % of the input variables (last line) and 7 PCs. Using the criterion defined above the maximum reduction of the variable input space was 30% ($P=70\%$). Using only 3 predictor variables a correlation of $R=0.862$ ($p<0.01$) could be achieved.

Table 1. Evaluation of the predictive quality for the prediction of peak-torque to breakaway

<i>PCs</i>	<i>R</i>	<i>RAE [%]</i>	<i>MAE[Nm]</i>	<i>exp</i>	<i>C</i>	<i>P [%]</i>
1	0.630	74	1.66	1	0.5	70
2	0.800	49	1.1	2	3	90
3	0.862	42	0.94	3	1	100
4	0.850	44	1.34	2	0.3	70
5	0.905	40.2	1.26	1	0.3	80
6	0.952	30	0.67	1	0.5	80
7	0.975	22.4	0.503	1	0.5	100

For test setting B a model of the whole proximal femur including the small trochanter was created. Analyzing the registration accuracy a mean absolute error of 9.1 grey levels ($\sigma=7.1$), which corresponds to 3.6% of the whole intensity range, was measured. The results of the leave-one-out tests are summed up in table 2, which is having the same structure as table 1. Using 100 % of the input variables (last line), the highest correlation $R=0.93$ ($p<0.01$) could be obtained with 10 predictor variables. The variable input space could be reduced up to 50%, which also lead to the highest correlation ($R=0.944$, $p<0.01$). Using 3 predictor variables still a correlation of $R=0.886$ ($p<0.01$) could be achieved. Note, that for both testing scenarios all results have been obtained using n-fold cross-validation

Table 2. Evaluation of the predictive quality for the prediction of failure load

<i>PCs</i>	<i>R</i>	<i>RAE [%]</i>	<i>MAE[N]</i>	<i>exponent</i>	<i>C</i>	<i>P [%]</i>
1	0.755	59	602	2	0.5	50
2	0.763	55	552	1	5	60
3	0.886	36	368	3	1	50
4	0.830	45.6	459	3	0.7	50
5	0.880	41	412	2	0.3	50
6	0.944	29	301	2	10	50
7	0.891	39.8	401	1	0.7	60
8	0.910	34.3	344	1	3	100
9	0.930	30.6	308	1	5	100
10	0.930	27.6	278	1	10	100

For this study DXA measurements in 5 different regions in the head and neck region of the proximal femur were also obtained. The DXA values measured in the femoral neck showed the highest correlation to failure load with $R=0.73$. Figure 1 is illustrating what structures are described by the 3 PCs showing highest correlation to failure load on one coronal sample slice.

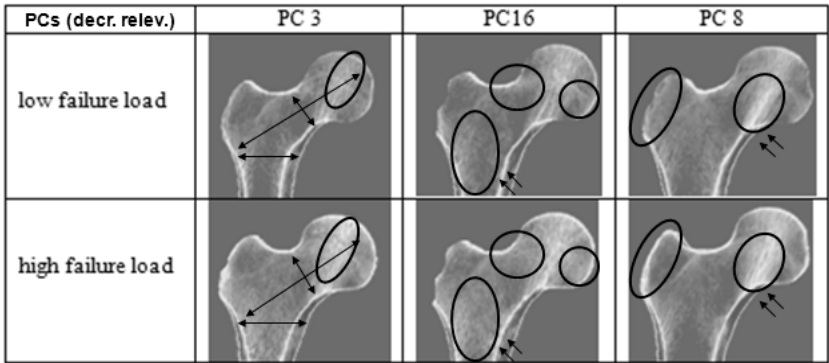


Fig. 1. Variations of the proximal femur described by the 3 Principal Components with highest correlation to failure load (Coronal cut through femoral head neck and shaft)

The images illustrate the major differences between femora with low (top) and high fracture load (bottom). PC 3 predominantly illustrates variations of the shape of the femur (neck thickness, shape of head, shape of greater trochanter) and the cancellous bone in the femoral head. The two other PCs mainly describe variations of the cancellous bone in the shaft, neck (PC 16) and head region (PCs 16 and 8). PC 8 also illustrates variations concerning the shape and length of the cortical region in the lower femoral neck. The regions around the greater and small trochanter, around the femoral head and neck as well as the cortical region in the shaft also remained in the model after 50 % reduction of the input variables.

4 Discussion

Using two different testing scenarios and an extensive cross-validation scheme it has been demonstrated that the model based approach presented in this work, can provide an excellent tool for estimating femoral fracture load and peak torque to breakaway with high accuracy using CT images with clinical resolution. By using this approach one has the possibility to depict a number of different geometric and structural properties that can be used to predict biomechanical bone parameters, accurately and reliably. Although the tests were performed using images with comparably low resolution, the quantitative results in terms of the amount of correlation between real and predicted parameter values are absolutely comparable to studies that partly use higher image resolutions [7-10].

Beside of the accuracy of the proposed method, another positive aspect of the algorithm is its ability to implicitly identify potential parameters as well as sub-regions that are useful for the prediction of biomechanical properties. The nature of the parameters that can be identified range from geometrical properties of the bone to variations of the spatial intensity distribution in medical images. In combination with the ease of use of the proposed algorithm this fact makes the approach useful for a wide field of applications. Having a valid (regression) model, the calculation of a shape-intensity representation and parameter prediction for a new unseen segmented dataset only takes ~5-15 minutes, depending on the resolution of the CT images. The next

steps in this project will be the investigation of alternative methods for dimensionality reduction and the application of the proposed methods in a large clinical study, which has been started recently.

References

1. Adams, J.E.: Single and dual energy X-ray absorptiometry. *Eur. Radiol.* 7, 20–31 (1997)
2. Faulkner, K.G., Cummings, S.R., Black, D., Palermo, L., Gluer, C.C., Genant, H.K.: Simple measurement of femoral geometry predicts hip fracture: the study of osteoporotic fractures. *J. Bone Miner. Res.* 8, 1211–1217 (1993)
3. Bergot, C., Bousson, V., Meunier, A., Laval-Jeantet, M., Laredo, J.D.: Hip fracture risk and proximal femur geometry from DXA scans. *Osteoporos Int.* 13, 542–550 (2002)
4. Boehm, H.F., Link, T., Monetti, R., Mueller, D., Rummeny, E., Newitt, D., Majumdar, S., Raeth, C.W.: Application of the Minkowski Functionals in 3D to High Resolution MR Images of Trabecular Bone for the Prediction of the Biomechanical Strength. *Microscopy and Microanalysis*, 716–717 (2004)
5. Link, T.M., Majumdar, S., Lin, J.C., Augat, P., Gould, R.G., Newitt, D., Ouyang, X., Lang, T.F., Mathur, A., Genant, H.K.: Assessment of trabecular structure using high resolution CT images and texture analysis. *J. Comput. Assist. Tomogr.* 22, 15–24 (1998)
6. Bessho, M., Ohnishi, I., Matsuyama, J., Matsumoto, T., Imai, K., Nakamura, K.: Prediction of strength and strain of the proximal femur by a CT-based finite element method. *J. Biomech.* 40, 1745–1753 (2007)
7. Keyak, J.H.: Improved prediction of proximal femoral fracture load using nonlinear finite element models. *Med. Eng. Phys.* 23, 165–173 (2001)
8. Suhr, N., Hengg, C., Schwyn, R., Windolf, M., Quarz, V., Hänni, M.: Mechanical torque measurement predicts load to implant cut-out: a biomechanical study investigating DHS anchorage in femoral heads. *Arch. Orthop. Trauma. Surg.* (2006) (online)
9. Fritscher, K.D., Schuler, B., Grünerbl, A., Hänni, M., Schwieger, K., Suhr, N., Schubert, R.: Assessment of femoral bone quality using co-occurrence matrices and adaptive regions of interest. In: *Proceedings of SPIE Medical Imaging*, vol. 6514, p. 65141K (2007)
10. Huber, M.B., Carballido-Gamio, J., Fritscher, K.D., Schubert, R., Haenni, M., Hengg, D., Majumdar, S., Link, T.M.: Morphological Texture Analysis of Radiographs of the Proximal Femur- in vitro study using biomechanical strength as a standard of reference. In: *29th ASBMR, Honolulu* (2007)
11. Rueckert, D., Frangi, A.F., Schnabel, J.A.: Automatic Construction of 3-D Statistical Deformation Models of the Brain Using Nonrigid Registration. *IEEE Trans. Med. Imaging* 22, 1014–1025 (2001)
12. Fritscher, K.D., Gruenerbl, A., Schubert, R.: 3D image segmentation using combined shape-intensity prior models. *International Journal of Radiology and Surgery* 1(3), 123–135 (2007)
13. Leventon, M., Grimson, E., Faugeras, O.: Statistical Shape Influence in Geodesic Active Contours. *Computer Vision and Pattern Recognition* 1, 316–323 (2000)
14. Hall, M.: Correlation-based Feature Selection for Machine Learning. Department of Computer Science. PhD Thesis. Waikato University, New Zealand (1998)
15. Smola, A.J., Schoelkopf, B.: A tutorial on support vector regression. *Statistics and Computing* 14, 199–222 (2003)
16. Lochmüller, E.M., Eckstein, F.: Biomechanische Tests in der Evaluation osteodensitometrischer Verfahren - Hintergrund, Übersicht und aktuelle Befunde. *Osteologie* 11, 154–177 (2002)