Medical Engineering and Physics 63 (2019) 88–92

Contents lists available at ScienceDirect

Medical Engineering and Physics

journal homepage: www.elsevier.com/locate/medengphy

Technical note

Efficacy and efficiency of multivariate linear regression for rapid prediction of femoral strain fields during activity

Hamed Ziaeipoor[a,*], Saulo Martelli[a,b], Marcus Pandy[c], Mark Taylor[a]

[a] Medical Device Research Institute, College of Science and Engineering, Flinders University, Clovelly Park, Tonsley, Adelaide, SA, Australia
[b] NorthWest Academic Centre, The University of Melbourne, St Albans, VIC, Australia
[c] Department of Mechanical Engineering, University of Melbourne, Parkville, VIC, Australia

ARTICLE INFO

Article history:
Received 23 January 2018
Revised 19 November 2018
Accepted 4 December 2018

Keywords:
Musculoskeletal
Finite-element
Surrogate model
Human gait

ABSTRACT

Multivariate Linear Regression-based (MLR) surrogate models were explored to reduce the computational cost of predicting femoral strains during normal activity in comparison with finite element analysis. The musculoskeletal model of one individual, the finite-element model of the right femur, and experimental force and motion data for normal walking, fast walking, stair ascent, stair descent, and rising from a chair were obtained from a previous study. Equivalent Von Mises strain was calculated for 1000 frames uniformly distributed across activities. MLR surrogate models were generated using training sets of 50, 100, 200 and 300 samples. The finite-element and MLR analyses were compared using linear regression. The Root Mean Square Error (RMSE) and the 95th percentile of the strain error distribution were used as indicators of average and peak error. The MLR model trained using 200 samples (RMSE < 108 $\mu\varepsilon$; peak error < 228 $\mu\varepsilon$) was used as a reference. The finite-element method required 66 s per frame on a standard desktop computer. The MLR model required 0.1 s per frame plus 1848 s of training time. RMSE ranged from 1.2% to 1.3% while peak error ranged from 2.2% to 3.6% of the maximum micro-strain (5020 $\mu\varepsilon$). Performance within an activity was lower during early and late stance, with RMSE of 4.1% and peak error of 8.6% of the maximum computed micro-strain. These results show that MLR surrogate models may be used to rapidly and accurately estimate strain fields in long bones during daily physical activity.

1. Introduction

Quantifying femoral strain distribution is important for studying bone adaptation [1–3], diagnosing individuals most at risk of femoral fracture [4–6], and optimizing the biomechanical behaviour of implantable devices [7,8]. Over the last few decades, finite-element analysis has been used extensively to quantify the entire femoral strain field [9–11], and there is growing interest in using this method to characterise strain distributions in multiple individuals [12,13] and across multiple trials and tasks [14]. In addition, there is need to investigate the influence of the musculoskeletal (MS) modelling process on femoral strain predictions, by performing probabilistic analyses to account for uncertainties in the MS model inputj parameters [14–16] and examining alternative muscle recruitment strategies [17]. Unfortunately, the computational cost of performing such analyses can be prohibitive, thus new methods are needed to accurately and rapidly estimate the femoral strain field to enable large-scale studies of 100′s–1000′s of simulations to be performed.

Surrogate models represent a viable solution in that they can be trained using finite element calculations of femoral strain for a limited number of training sets and then used to rapidly provide femoral strain estimates for an arbitrary frame of motion or an entire activity. A variety of surrogate models have been used by the biomechanics community including Multivariate Linear Regression [18,19], Bayesian modelling [20], Artificial Neural Networks [18,21,22], Random Forest [23] and Kriging [24–26], either for linear problems, (e.g., assessment of femoral neck fracture during a single load case [18]) or for non-linear problems (e.g., modelling the contact between bone and implant [19]). Most studies predict a single scalar outcome, such as joint moments and muscle forces [27]; contact forces and contact pressure [21,25,26,28–30]; femoral neck strain and fracture load [18]; implant micro-movement and stress shielding [20,31]. Multivariate Linear Regression has been used for predicting femoral neck strain [32], fracture load [18] and the micro-movement at the bone-implant interface [19]. However, the error and computational advantage of MLR over finite-element models remains unclear for the calculation of strain over the femoral volume and across normal activities of daily living.

* Corresponding author.
E-mail addresses: hamed.ziaeipoor@flinders.edu.au, ziae0005@flinders.edu.au (H. Ziaeipoor).

https://doi.org/10.1016/j.medengphy.2018.12.001
1350-4533/

The aim of this work was to explore the use of MLR for predicting femoral strain fields for a range of activities of daily living. Muscle forces, joint reaction forces and femoral strain were calculated for a single individual performing five tasks using a previously developed musculoskeletal and finite-element model [16]. A MLR surrogate model was trained using femoral strain, muscle forces, and joint reaction forces for a limited number of randomly selected frames of motion and then used to estimate femoral strain for multiple motor tasks. Model performance was assessed by comparing MLR estimates of the femoral strain field to corresponding results obtained from finite-element calculations.

2. Materials and methods

2.1. Data

A full-body musculoskeletal model, finite-element model of the femur of the dominant leg, marker trajectories, and ground reaction forces for a single healthy participant (female, 68-year-old, 53 kg weight, 157 cm height) were obtained from a previous study [16]. All experimental and computational methods are described in detail by Martelli et al. [16] and Dorn et al. [33], respectively. Briefly, marker trajectories and ground reaction forces were recorded for five trials of each of the following five tasks: walking at the self-selected speed (normal walking), fast walking, stair ascent, stair descent, and rising from and sitting down on a chair (chair rise). Trials with incomplete marker trajectories were discarded, resulting in five trials each for normal walking, fast walking and stair descent; four trials for stair ascent; and one trial for chair rise. A participant-specific musculoskeletal model was created by scaling the generic model described by Dorn et al. [33] using the segment lengths and body mass measured during a static trial.

The marker trajectories were labelled using a VICON motion capture system (Vicon, Oxford, UK), saved as c3d files, and then converted into OpenSim format using MOtoNMS [34]. Joint angles, muscle forces, and joint reaction forces were calculated using, respectively, inverse kinematics, static optimization and joint reaction analysis tools available in OpenSim [35]. The finite-element model of the right femur was a locally-isotropic, linear-elastic model whose geometry and element-by-element material properties were extracted from calibrated computed-tomography images following a well-established procedure [36]. The finite-element model was loaded by applying muscle forces and the hip joint reaction force for 50 frames uniformly distributed over each activity. The FE model was kinematically constrained distally (Fig. 1). The musculoskeletal and finite-element models were coupled using custom software [16]. The equivalent von Mises strain was calculated at each element centroid for a total of 1000 frames (20 trials, 50 frames per trial) as a compact indicator of both compressive and tensile strain states.

2.2. Multi-variate linear regression surrogate model

A Latin Hypercube (LH) sampling method was used to create the training set, which comprised of muscle forces, joint reaction forces and femoral strains for randomly selected frames of motion (Fig. 1). The process was repeated to generate four training sets consisting of 50, 100, 200 and 300 frames, respectively. Training sets of similar size have been used to develop surrogate models in previous studies [19,24]. The surrogate model, relating the applied forces to the equivalent von Mises strain, was developed by fitting a MLR model for each element. The model took the form:

$$\varepsilon^{j,k} = \sum_{i=0}^{25} \left(c_i \times f_i^k \right); f_i^k, \ldots, f_{25}^k$$

where $\varepsilon^{j,k}$ is the equivalent von Mises strain at element j for frame k, and c_i is the coefficient for the force i at frame k. The total number of forces applied to the finite-element model was 25, which included all the muscle forces in the musculoskeletal model acting on the femur and the hip reaction force. The strain field for all 1000 frames of motion was calculated using the calculated coefficients c_i in the MLR model and corresponding muscle and joint reaction forces. Performance of the MLR surrogate models was assessed by calculating the coefficient of determination (R^2) and the slope of the linear regression between the strains predicted by the surrogate and finite-element models. CPU times needed to complete the finite-element analysis, train the MLR models, and calculate femoral strain using the MLR models were compared on a standard desktop computer (8 CPUs Intel® Core(TM)® 3.4 GHz processor, 32 GB RAM). Strain error was calculated using the finite-element strain as a reference and evaluated using the Root Mean Square Error (RMSE) as well as the 95th percentile of the strain error distribution as an indicator of peak error. These parameters were analysed frame-by-frame within each trial (i.e. $RMSE_{Frame}$; R^2_{Frame}) by amalgamating all frames for each trial (i.e. $RMSE_{Trial}$; R^2_{Trial}) and for each activity (i.e. $RMSE_A$; R^2_A).

3. Results

The trial-by-trial comparison showed that the coefficient of determination and slope were close to unity for the training datasets greater than 50 ($R^2_{Trial} = 0.84 - 0.94$; $slope_{Trial}$=0.97−0.99). The prediction error of the surrogate model was a function of the size of the training set. Increasing the size of the training set from 100 to 200 frames reduced the average RMSE across trials from 132 $\mu\varepsilon$ to 108 $\mu\varepsilon$, while a relatively small decrease in RMSE to 107 $\mu\varepsilon$ was obtained by increasing the training set size to 300 samples (Table 1). Based on these observations, the remainder of the results are presented only for the MLR model trained using 200 samples.

CPU time for predicting the full femoral strain for all 1000 frames was 66,000 s using the finite-element model alone (i.e., 55 min were necessary for predicting femoral strain for an entire activity of 50 frames). Training the MLR model required 13,200 s for completing the 200 finite-element simulations in the training set, 528 s for training and 100 s for predicting all 1000 frames, which corresponds to 5 s for predicting femoral strain for an entire activity (50 frames). The MLR-based surrogate model was faster than finite-element analysis for solving 209 frames or more (Fig. 2).

Similar performance of the MLR model was observed for all activities. The median $RMSE_A$ varied between 80 $\mu\varepsilon$ for normal walking and 124 $\mu\varepsilon$ for chair rise. Peak $RMSE_A$ varied from 163 $\mu\varepsilon$ for stair ascent to 389 $\mu\varepsilon$ for chair rise (Fig. 3).

The performance of the MLR model is presented for a selected trial of normal walking as an exemplar activity (Figs. 4 and 5). Close visual agreement was observed between the strain distributions estimated by the surrogate model and those predicted by the FE model (Fig. 4). The average RMSE and peak error were 78 and 181 $\mu\varepsilon$, respectively, across different frames. RMSE reached 207 $\mu\varepsilon$ during early stance and 140 $\mu\varepsilon$ during late stance while the corresponding peak errors reached 433 $\mu\varepsilon$ and 391 $\mu\varepsilon$, respectively, for early and late stance (Fig. 5). The peak error was 8.6% of peak equivalent strain in the diaphysis, ranging from approximately 2920 to 5020 $\mu\varepsilon$ during the stance phase of gait. The average coefficient of determination and slope were 0.97 and 0.99, respectively.

4. Discussion

Finite element analysis has been used extensively in orthopaedic biomechanics research [37], but there are a number of

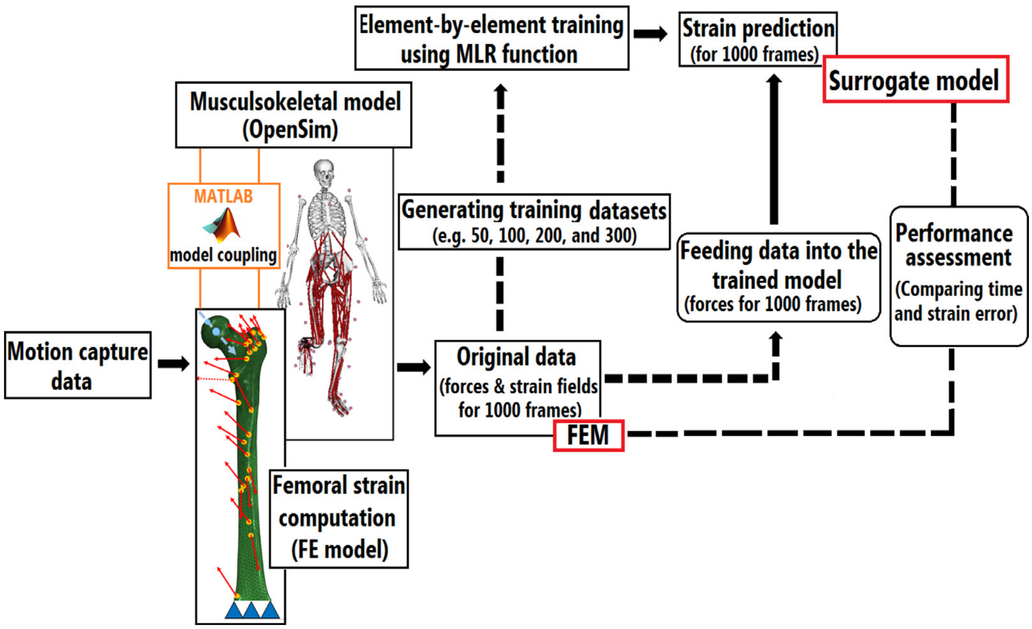

Fig. 1. Flowchart illustrating the linear-based surrogate modelling approach used in the present study.

Table 1
Effect of the size of training datasets on the accuracy of model-predicted femoral strains. Model accuracy was evaluated by computing the mean and peak error and the mean of coefficient of determination. These reported errors are based on pooled data.

Training Datasets	Mean RMSE ($\mu\varepsilon$)	Peak RMSE ($\mu\varepsilon$)	Mean R^2	Training Time (min)
50	227	408,484	0.84	8.5
100	132	326	0.92	8.7
200	108	228	0.94	8.8
300	107	201	0.94	8.9

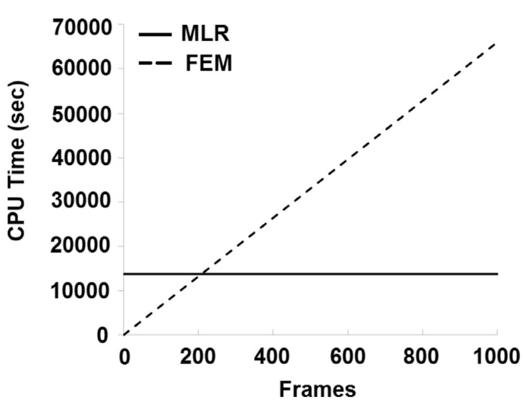

Fig. 2. CPU time required by the finite-element model and MLR model plotted against the number of frames.

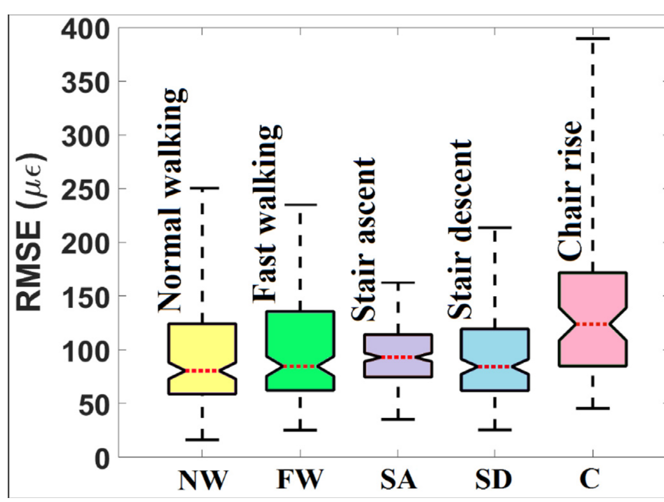

Fig. 3. Box plots used to quantify the accuracy of model-predicted strains obtained from MLR surrogate modelling. The black box represents the range of the error between the 25th and 75th percentiles while the red horizontal dashed line represents the median error. The black dashed line represents the 95th percentile of $RMSE_A$ for each activity.

barriers involved in the translation of FE modelling to the clinic. One problem is that predicting the full femoral strain for multiple tasks using a coupled FE-musculoskeletal modelling approach is computationally expensive. The current study represents a first step in overcoming this barrier, by demonstrating that reliable estimates of strain distributions may be obtained rapidly. Surrogate models offer a potentially powerful alternative as they provide predictions of bone strains in seconds rather than hours. The present study evaluated the performance of a multivariate linear regression surrogate model in approximating the full strain field of an intact femur during five different activities of daily living.

We found that reliable predictions of femoral strain could be obtained across all five activities by training the surrogate model using 200 samples. The surrogate model closely reproduced the FE results at a low computational cost, with typical solution times of 5 s per activity (50 frames) compared to 55 min needed for a finite-element analysis.

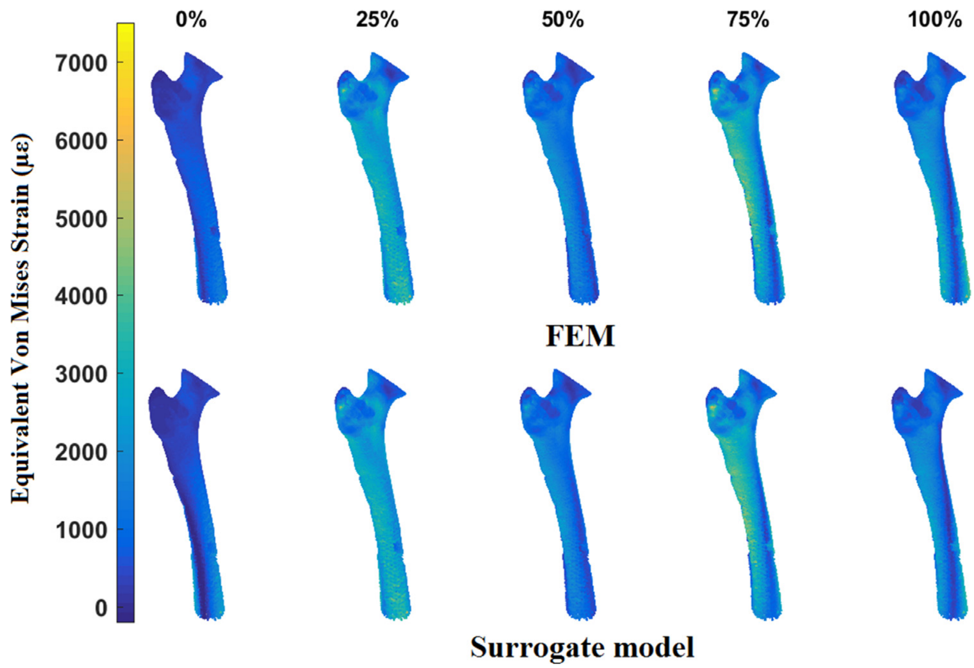

Fig. 4. Contour plots showing the calculated femoral strain fields for normal walking obtained by applying finite element modelling (FEM) and MLR surrogate modelling. Results are shown at 25% intervals of the stance phase. 0% and 100% indicate the stance phase.

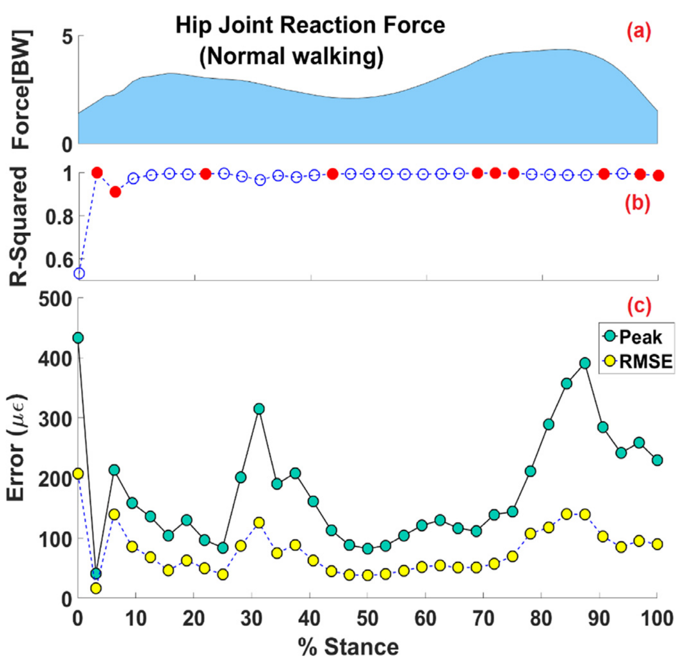

Fig. 5. Evaluating the performance of the MLR surrogate model for normal walking: (a) pattern of the hip joint reaction force; (b) coefficient of determination (R^2_{Frame}); (c) peak error and root mean square error ($RMSE_{Frame}$) at each frame. BW in part (a) refers to body weight; the red dots shown in part (b) represent the frames used to train the surrogate model.

The predicted strains from the MLR model were in close agreement with those obtained using the finite-element model. The peak error in the MLR model was 8.6% of the peak equivalent strain (5020 με), which is comparable to the error (i.e., 4.2–8.3% of peak strain on average) caused by material properties and geometry errors committed while generating the finite-element models from calibrated computed-tomography images [38]. Furthermore, the average RMSE was 78 με, which is consistent with the average error (113 $\mu\varepsilon$) obtained when finite-element models are used to predict experimentally measured cortical strains [38]. Therefore, MLR models represent valid surrogates of finite-element calculations of femoral strain during activity. The training sample size was similar to that reported in previous surrogate modelling studies in biomechanics: Fitzpatrick et al. (2014) required 100–200 samples for a MLR-based surrogate model; Taylor et al. (2017) needed 200–500 samples to train an artificial neural network; and Lin et al. (2009) required 300 samples to develop a kriging-based surrogate model. This supports the validity of the MLR model developed in the present study.

The current study is not without limitations. Firstly, Latin Hypercube sampling was used to generate the training datasets, but generating more uniformly distributed samples using other potential techniques may improve model accuracy. Secondly, the performance of the surrogate model was lowest during early stance where the coefficient of determination was only 0.53. This error is likely caused by the non-linear behaviour of the model, arising, for example, from the displacement of the hip centre of pressure during motion. Different surrogate methods (e.g. MARS, Gaussian Process and Artificial Neural Networks) may further improve model performance. Thirdly, the prediction time of the MLR model (0.1 s per frame) was much faster than that of the finite-element model, although the MLR required 200 finite-element simulations for generating the training set and 528 s for training the model. Thus, the MRL model is computationally advantageous relative to the finite-element model only when 209 frames of motion or more are to be analysed (Fig. 2). Fourth, only normal activities were included in the reference study [16] to limit the risk of injury for the participants while executing demanding (e.g., sprinting) or para-physiological (e.g., falling) activities. Therefore, the validity of the present conclusion is limited to normal locomotion. Fifth, the MLR model was developed for a single healthy individual possibly limiting the generality of the present conclusions. However, the strain range predicted by the model (0–5020 $\mu\varepsilon$) spans a large portion of physiologically admissible strains [39] and the loading conditions did span a broad range of normal activities, providing confidence on the performance of the MLR model over a relevant range of femoral strain and boundary conditions of the femur.

5. Conclusions

A Multivariate Linear Regression model was successfully developed for a single individual and used to rapidly predict the full femoral strain field for a range of activities of daily living. The MLR model was able to predict the femoral strain field for each studied activity within an error comparable to the intrinsic error in finite-element models based on clinical CT images and was computationally advantageous when 209 loading cases or more were analysed. Hence, MLR enables large statistical studies of femoral strain during activity.

Competing interests

There are no conflicts of interest associated with the work performed in this study.

Funding

This work was supported by the Australian Government Research Training Program Scholarship (AGRTPS); and the Australian Research Council (Grant no. DP180103146).

Ethical approval

Not required.

References

[1] Geraldes DM, Modenese L, Phillips ATM. Consideration of multiple load cases is critical in modelling orthotropic bone adaptation in the femur. Biomech Model Mechanobiol 2016;15:1029–42.
[2] Adams DJ, Spirt AA, Brown TD, Fritton SP, Rubin CT, Brand RA. Testing the daily stress stimulus theory of bone adaptation with natural and experimentally controlled strain histories. J Biomech 1997;30:671–8.
[3] Duda GN, Heller M, Albinger J, Schulz O, Schneider E, Claes L. Influence of muscle forces on femoral strain distribution. J Biomech 1998;31:841–6.
[4] Bessho M, Ohnishi I, Matsumoto T, Ohashi S, Matsuyama J, Tobita K, et al. Prediction of proximal femur strength using a CT-based nonlinear finite element method: differences in predicted fracture load and site with changing load and boundary conditions. Bone 2009;45:226–31.
[5] 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 2007;40:1745–53.
[6] Lotz JC, Cheal EJ, Hayes WC. Fracture prediction for the proximal femur using finite element models: part II–nonlinear analysis. J Biomech Eng 1991;113:361–5.
[7] Helwig P, Faust G, Hindenlang U, Hirschmuller A, Konstantinidis L, Bahrs C, et al. Finite element analysis of four different implants inserted in different positions to stabilize an idealized trochanteric femoral fracture. Injury 2009;40:288–95.
[8] Chanda S, Dickinson A, Gupta S, Browne M. Full-field in vitro measurements and in silico predictions of strain shielding in the implanted femur after total hip arthroplasty. Proc Inst Mech Eng H 2015;229:549–59.
[9] Martelli S, Pivonka P, Ebeling PR. Femoral shaft strains during daily activities: implications for atypical femoral fractures. Clin Biomech 2014;29:869–76.
[10] Martelli S. Femoral neck strain during maximal contraction of Isolated hip-spanning muscle groups. Comput Math Methods Med 2017 Article ID 2873789. doi:10.1155/2017/2873789.
[11] Al-Dirini RMA, Huff D, Zhang J, Besier T, Clement JG, Taylor M. Influence of collars on the primary stability of cementless femoral stems: a finite element study using a diverse patient cohort. J Orthop Res 2018;36:1185–95.
[12] Bryan R, Nair PB, Taylor M. Use of a statistical model of the whole femur in a large scale, multi-model study of femoral neck fracture risk. J Biomech 2009;42:2171–6.
[13] Zhang J, Besier TF. Accuracy of femur reconstruction from sparse geometric data using a statistical shape model. Comput Methods Biomech Biomed Eng 2016;20:566–76.
[14] Martelli S, Valente G, Viceconti M, Taddei F. Sensitivity of a subject-specific musculoskeletal model to the uncertainties on the joint axes location. Comput Methods Biomech Biomed Eng 2015;18:1555–63.
[15] Ackland DC, Lin YC, Pandy MG. Sensitivity of model predictions of muscle function to changes in moment arms and muscle-tendon properties: a Monte-Carlo analysis. J Biomech 2012;45:1463–71.
[16] Martelli S, Kersh ME, Pandy MG. Sensitivity of femoral strain calculations to anatomical scaling errors in musculoskeletal models of movement. J Biomech 2015;48:3606–15.
[17] Martelli S, Calvetti D, Somersalo E, Viceconti M. Stochastic modelling of muscle recruitment during activity. Interface Focus 2015;5:20140094. doi:10.1098/rsfs.2014.0094.
[18] Taylor M, Perilli E, Martelli S. Development of a surrogate model based on patient weight, bone mass and geometry to predict femoral neck strains and fracture loads. J Biomech 2017;55:121–7.
[19] Fitzpatrick CK, Hemelaar P, Taylor M. Computationally efficient prediction of bone-implant interface micromotion of a cementless tibial tray during gait. J Biomech 2014;47:1718–26.
[20] Bah MT, Nair PB, Taylor M, Browne M. Efficient computational method for assessing the effects of implant positioning in cementless total hip replacements. J Biomech 2011;44:1417–22.
[21] Eskinazi I, Fregly BJ. Surrogate modeling of deformable joint contact using artificial neural networks. Med Eng Phys 2015;37:885–91.
[22] Liang L, Liu M, Martin C, Sun W. A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis. J R Soc Interface 2018;15:20170844. doi:10.1098/rsif.2017.0844.
[23] Donaldson FE, Nyman E, Coburn JC. Prediction of contact mechanics in metal-on-metal total hip replacement for parametrically comprehensive designs and loads. J Biomech 2015;48:1828–35.
[24] O'Rourke D, Martelli S, Bottema M, Taylor M. A computational efficient method to assess the sensitivity of finite-element models: an illustration with the hemipelvis. J Biomech Eng 2016;138. doi:10.1115/1.4034831.
[25] Walter JP, Pandy MG. Dynamic simulation of knee-joint loading during gait using force-feedback control and surrogate contact modelling. Med Eng Phys 2017;48:196–205.
[26] Lin YC, Haftka RT, Queipo NV, Fregly BJ. Two-dimensional surrogate contact modeling for computationally efficient dynamic simulation of total knee replacements. J Biomech Eng 2009;131:041010. doi:10.1115/1.3005152.
[27] Favre J, Hayoz M, Erhart-Hledik JC, Andriacchi TP. A neural network model to predict knee adduction moment during walking based on ground reaction force and anthropometric measurements. J Biomech 2012;45:692–8.
[28] Pizzolato C, Reggiani M, Saxby DJ, Ceseracciu E, Modenese L, Lloyd DG. Biofeedback for gait retraining based on real-time estimation of tibiofemoral joint contact forces. IEEE Trans Neural Syst Rehabil Eng 2017;25:1612–21.
[29] Lin YC, Haftka RT, Queipo NV, Fregly BJ. Surrogate articular contact models for computationally efficient multibody dynamic simulations. Med Eng Phys 2010;32:584–94.
[30] Halloran JP, Erdemir A, van den Bogert AJ. Adaptive surrogate modeling for efficient coupling of musculoskeletal control and tissue deformation models. J Biomech Eng 2009;131:011014. doi:10.1115/1.3005333.
[31] Cilla M, Borgiani E, Martínez J, Duda GN, Checa S. Machine learning techniques for the optimization of joint replacements: application to a short-stem hip implant. PLoS One 2017;12:e0183755. doi:10.1371/journal.pone.0183755.
[32] Kersh ME, Martelli S, Zebaze R, Seeman E, Pandy MG. Mechanical loading of the femoral neck in human locomotion. J Bone Miner Res 2018;33:1999–2006. doi:10.1002/jbmr.3529.
[33] Dorn TW, Schache AG, Pandy MG. Muscular strategy shift in human running: dependence of running speed on hip and ankle muscle performance. J Exp Biol 2012;215:1944–56.
[34] Mantoan A, Pizzolato C, Sartori M, Sawacha Z, Cobelli C, Reggiani M. MOtoNMS: a MATLAB toolbox to process motion data for neuromusculoskeletal modeling and simulation. Source Code Biol Med 2015;10:1–14.
[35] Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, et al. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Trans Biomed Eng 2007;54:1940–50.
[36] Taddei F, Cristofolini L, Martelli S, Gill HS, Viceconti M. Subject-specific finite element models of long bones: an in vitro evaluation of the overall accuracy. J Biomech 2006;39:2457–67.
[37] Taylor M, Prendergast PJ. Four decades of finite element analysis of orthopaedic devices: where are we now and what are the opportunities? J Biomech 2015;48:767–78.
[38] Taddei F, Martelli S, Reggiani B, Cristofolini L, Viceconti M. Finite-element modeling of bones from CT data: sensitivity to geometry and material uncertainties. IEEE Trans Biomed Eng 2006;53:2194–200.
[39] Bayraktar HH, Morgan EF, Niebur GL, Morris GE, Wong EK, Keaveny TM. Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. J Biomech 2004;37:27–35.