Flinders UNIVERSITY

SENSITIVITY OF FEMORAL STRAIN CALCULATIONS TO ANATOMICAL SCALING ERRORS IN MUSCULOSKELETAL MODELS OF MOVEMENT

Saulo Martelli[1,2], Mariana E. Kersh[3], Marcus G. Pandy[4]

[1]Medical Device Research Institute, School of Computer Science, Engineering and Mathematics, Flinders University, Bedford Park, Australia

[2]North West Academic Centre, The University of Melbourne, St Albans, Australia

[3]Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Illinois, USA

[4]Department of Mechanical Engineering, University of Melbourne, Parkville, Australia

Submitted as an original full-length article to the Journal of Biomechanics

3 August 2015

Word count: Text (Introduction to Discussion): 3,797; Abstract: 250

Corresponding author:

Saulo Martelli, Ph.D.

Medical Device Research Institute

School of Computer Science, Engineering and Mathematics

Flinders University

Sturt Rd, Bedford Park SA 5042, Australia

E-mail: saulo.martelli@flinders.edu.au

ABSTRACT

The determination of femoral strain in post-menopausal women is important for studying bone fragility. Femoral strain can be calculated using a reference musculoskeletal model scaled to participant anatomies (referred to as scaled-generic) combined with finite-element models. However, anthropometric errors committed while scaling affect the calculation of femoral strains. We assessed the sensitivity of femoral strain calculations to scaled-generic anthropometric errors. We obtained CT images of the pelves and femora of ten healthy post-menopausal women and collected gait data from each participant during six weight-bearing tasks. Scaled-generic musculoskeletal models were generated using skin-mounted marker distances. Image-based models were created by modifying the scaled-generic models using muscle and joint parameters obtained from the CT data. Scaled-generic and image-based muscle and hip joint forces were determined by optimization. A finite-element model of each femur was generated from the CT images, and both image-based and scaled-generic principal strains were computed in 32 regions throughout the femur. The intra-participant regional RMS error increased from 380 με (R^2=0.92, $p<0.001$) to 4,064 με (R^2=0.48, $p<0.001$), representing 5.2% and 55.6% of the tensile yield strain in bone, respectively. The peak strain difference increased from 2,821 με in the proximal region to 34,166 με at the distal end of the femur. The inter-participant RMS error throughout the 32 femoral regions was 430 με (R^2=0.95, $p<0.001$), representing 5.9% of bone tensile yield strain. We conclude that scaled-generic models can be used for determining cohort-based averages of femoral strain whereas image-based models are better suited for calculating participant-specific strains throughout the femur.

Keywords: Finite-element femur model; Scaled-generic; Image-based musculoskeletal model; Subject-specific bone strain; Anatomical scaling.

1. Introduction

The quantification of femoral strain during daily activities is important for understanding the biomechanical implications of osteoporosis (van Rietbergen et al., 2003), for which post-menopausal women are most at risk. For example, intra-participant femoral strains can provide information about fracture risk (Cody et al., 1999) while inter-participant averages can provide insights into understanding the bone response to exercise treatments (Lang et al., 2014). In vivo femoral strains can be estimated non-invasively using a scaled-generic musculoskeletal model scaled to participant anatomies (herein referred to as 'scaled-generic models') combined with a finite-element model of the femur (Jonkers et al., 2008; Martelli et al., 2014a). However, errors in the definition of the model anthropometry affect calculation of muscle forces (Lenaerts et al., 2009), which likely propagate to bone strain calculation. Several studies have investigated the sensitivity of muscle and joint force calculations to uncertainties in anatomical and muscle parameters (Ackland et al., 2012; Correa et al., 2011; Martelli et al., 2015; Redl et al., 2007; Scheys et al., 2009; Xiao and Higginson, 2010) while others have examined the sensitivity of femoral strain calculations to uncertainties in measurements of the geometry and material properties of the femur (Taddei et al., 2006). To date, no study has investigated the sensitivity of femoral strain calculations to anthropometric errors arising from uncertainties in, for example, body-segmental masses and lengths.

Magnetic-resonance (MR) and computed-tomography (CT) images can provide detailed anthropometric information about the human musculoskeletal system. While MR imaging is the preferred method for acquiring muscle-tendon attachment sites and paths, joint centre positions, and the orientations of joint rotation axes (Blemker et al., 2007; Scheys et al., 2008), this approach is not suitable for extracting bone mineral density (BMD), which is needed to model the elastic properties of bone (Schileo et al., 2007). Alternatively, bone surfaces, joint centres and orientations can be determined by segmenting CT images (Taddei

et al., 2012), and the images' Housfield unit data can be used to describe the BMD and elastic property distributions (Schileo et al., 2007). Although the low contrast of CT images complicates extracting soft-tissue anatomical structures such as muscles, CT images can serve as a reference for registering a muscular system atlas to a participant's anatomy (Abdel Fatah et al., 2012; Taddei et al., 2012). Therefore, CT images can provide all information necessary to generate both musculoskeletal and finite-element models of a specific participant (herein referred to as 'image-based models').

Scaling procedures have been used to generate musculoskeletal models of participants by applying a limited number of anthropometric parameters to a scaling algorithm (Delp et al., 2007, 1990). Typically, the body mass and segment lengths in a generic-reference model are scaled to an individual participant using information from the skin-mounted marker positions and ground reaction forces acquired during a static pose, thereby creating a 'scaled-generic' model. Scaled-generic models have been successfully used to study general patterns of human motion (Correa et al., 2010; Delp et al., 1990). However, scaling causes unavoidable anthropometric errors, which in turn may compromise the assessment of individual features in muscle and joint force patterns (Lenaerts et al., 2009).

Previous studies addressing the sensitivity of scaled-generic models investigated different model outputs and reached different conclusions. Correa et al. (2011) concluded that scaled-generic models are as accurate as image-based models when evaluating the potential (per-unit-force) contributions of individual muscles to joint and centre-of-mass accelerations during walking. Lenaerts et al. (2009) concluded that participant-specific hip geometry is important in the calculation of hip contact forces while walking; they reported average differences between scaled-generic and image-based models of 0.52 times body weight (BW). No study has reported the sensitivity of femoral strain calculations to anthropometric errors committed while scaling a scaled-generic model to participants' anatomies. However,

this information is essential for understanding the limits of applicability of the model results (Viceconti et al., 2005).

The aim of this study was to investigate how anthropometric errors introduced when scaling a scaled-generic musculoskeletal model to a participant's anatomy propagate to femoral strain calculations. Femoral strains were computed using scaled-generic and image-based models of ten participants for six weight-bearing tasks. The influence of scaled-generic anthropometric errors was assessed by analysing a) participant-specific (intra-participant) femoral strains, and b) average (inter-participant) femoral strains within a cohort.

2. Materials and Methods

Ten healthy post-menopausal women (age, 66.7 ± 7.0 years; height, 159 ± 6.6 cm; weight, 66.3 ± 22.5 kg) were recruited to this study (Table 1). All participants could walk unassisted and had no reported history of musculoskeletal disease. Ethics approval for the study was obtained from the Human Research Ethics Committee at the University of Melbourne.

2.1. Data collection

CT images of each participant were obtained of the pelvic and thigh regions using a clinical whole-body scanner (Aquilon CT, Toshiba Corporation, Tokyo) and an axial scanning protocol (tube voltage: 120 kV; tube current: 200 mA). For each scan, two datasets of monochromatic, 16-bit, 512×512 pixel images with slice thickness of 0.5 mm and spacing of 0.5 mm were obtained. The femur dataset was reconstructed using an in-plane transverse resolution of 0.5×0.5 mm whereas the pelvis dataset was reconstructed using an adjusted in-plane transverse resolution to accommodate the entire pelvis. A five-sample (hydroxyapatite density range: 0-200 mg/cm^3) calibration phantom (Mindways Software, Inc., Austin, TX) was placed below the participant's dominant leg while scanning.

Gait analysis experiments were performed at the Biomotion Laboratory, University of Melbourne. Forty-six skin-mounted reflective markers were attached to anatomical locations as described by Dorn et al. (2012), including the pelvis (3), thigh (6), shank (5) and foot (6). The remaining markers were placed along the upper extremities and torso. Marker trajectories were recorded with a 10-camera motion capture system (VICON, Oxford Metrics Group, Oxford) sampling at 120 Hz. Each participant was instructed to (a) walk at a self-selected speed; (b) walk at a faster self-selected speed; (c) ascend and descend a flight of 3 steps (step height = 16.5 cm) at self-selected speeds while engaging with the first step of the staircase using the dominant foot; (d) rise from and sit on a chair (chair height = 47 cm); and (e) jump as high as possible from a comfortable standing position with each foot placed on a separate force platform. Five repetitions of each task were executed. Ground reaction forces and moments were recorded using three strain-gauged force plates (AMTI, Watertown, MA) sampling at 2000 Hz. The ground force data were low-pass filtered using a fourth-order, recursive, zero-lag, Butterworth filter with a cut-off frequency of 40 Hz. A static trial was recorded to measure the inter-marker distances. Marker trajectories were low-pass filtered using a second-order recursive, zero-lag, Butterworth filter with a cut-off frequency of 6 Hz.

2.2. Musculoskeletal modelling

The scaled-generic and image-based musculoskeletal models were based on the generic model developed by Dorn et al. (2012) . The generic model was comprised of 12 segments with 31 independent degrees-of-freedom actuated by 92 Hill-type muscle–tendon units (Fig. 1A). A ball-and-socket joint represented the lumbar joint, each shoulder, and each hip; a translating hinge joint represented each knee; and a universal joint represented each ankle. The shoulder and elbow joints were actuated by 10 ideal torque motors, while all other joints were actuated by Hill-type muscle–tendon units.

Scaled-generic models were obtained by scaling the generic model to match each participant's body anthropometry and mass using OpenSim (Delp et al., 2007). Inter-marker distances recorded during the static trial (Fig. 1B) were used to scale bone geometries, joint centres, joint rotation axes, muscle paths, fibre lengths, and tendon slack lengths. The mass of the generic model was scaled to match that of each participant by preserving the mass ratio between segments in the generic model. Image-based models were created using anthropometric measurements obtained from the CT images for the pelvis and femur segments, skin-marker locations for the torso, and scaled-generic parameters for the remaining segments. The geometries of the pelves and femora were segmented from the CT data using Amira (Visage Imaging GmbH, Burlington, MA). The hip joint centre was defined as the centre of the sphere used to best-fit the femoral head surface. The knee axis was assumed to be the axis connecting the femoral epicondyles, and the lumbar joint was assumed to be located at the antero-posterior level of the vertebral foramen and at the mid-point of the L5-S1 inter-vertebral space as identified in the sagittal plane. The torso was adjusted to match the vertical distance between the sacrum and the seventh cervical spine calculated from the skin-mounted markers (Fig. 1). Muscle paths in the scaled-generic model were registered on the skeletal surfaces by superimposing the muscle lines-of-action onto the CT data (Fig. 1C). The values of optimum muscle-fibre length and tendon slack length reported by Delp et al. (1990) were uniformly scaled so that each muscle developed its peak isometric force at the same joint angle in both the scaled-generic and image-based models.

Scaled-generic and image-based muscle and joint forces were calculated for the dominant leg of a selected trial. Joint angles were computed by performing an inverse kinematics analysis according to methods described by Delp et al. (2007). The joint angles and the measured ground reaction forces were used to calculate the net moment developed about each joint. Static optimisation was then used to decompose the net joint moments into muscle

forces by minimising the weighted sum of the squares of muscle activations (Anderson and Pandy, 2001). The hip joint force was calculated by solving for static equilibrium at the femur.

2.3 *Finite-element modelling*

Bone tissue was modelled using 10-node tetrahedral elements. A linear regression equation relating the grey levels in the CT data to the hydroxyapatite density contained in the five-sample calibration phantom was used to convert the images' grey levels into apparent bone density levels. The apparent bone density distribution was converted into an isotropic Young's modulus for each voxel using the relationships derived in Morgan et al. (2003). The Young's modulus values were integrated over each mesh element using Bonemat© (Super Computing Solutions, Bologna). The femur was partitioned into eight different levels: four diaphyseal, one pertrochanteric, and three femoral neck levels. Each level was further subdivided into four regions: anterior, posterior, medial and lateral aspects, giving 32 sub-regions altogether (Fig. 2). Each femur finite-element model was kinematically constrained at the femoral epicondyles, a condition that is statically equivalent to applying forces acting on the most distal femur (Martelli et al., 2014a). Five element layers surrounding the muscle attachment points were excluded to avoid boundary condition artefacts.

Scaled-generic and image-based muscle and hip joint forces were applied to the finite-element model using custom code developed in Matlab (MathWorks Inc., Natick, MA). The pelvic, femoral and tibial anatomical coordinate systems were calculated according to International Society of Biomechanics standards (Wu et al., 2002). The unit vector describing the line-of-action of each muscle force was assumed to originate at the muscle's attachment point on the femur and was oriented along the line-of-action of the muscle force. The muscle force components were obtained by multiplying the magnitude of the muscle force calculated

from static optimization by the unit force vector. The muscle force components were then applied at the node closest to the muscle attachment point in the finite-element model.

The hip joint force was applied to the node on the surface of the femoral head closest to the intersection between the hip contact force vector passing through the hip centre and the femoral head surface. Linear static simulations were performed in Abaqus© (Dassault Systemes, Vélizy-Villacoublay) using the implicit direct solver. The 90th percentile of the scaled-generic and image-based principal tensile and compressive strain values were calculated for each femoral sub-region over the course of 20 time steps during the load-bearing phase of each activity.

2.4 Metrics for comparing scaled-generic and image-based models

Image-based joint angles, joint moments, hip-joint contact forces, muscle activation patterns and femoral strains were compared with corresponding published values (Aamodt et al., 1997; Bergmann et al., 2001; Inman et al., 1989; Kadaba et al., 1989).

Anthropometric errors were defined as the difference between the scaled-generic and image-based joint-to-joint distances, femoral anteversion angles, caput-collum-diaphyseal (CCD) angles, femoral neck lengths, and muscle moment arms. Scaled-generic and image-based muscle and joint forces were compared using linear regressions. The moment generated by the image-based and scaled-generic force systems about six locations uniformly distributed between the mean constrained node at the distal femur and the hip joint centre was calculated. The distribution of the scaled-generic and image-based moment differences was assessed at each location.

The effect of scaled-generic anthropometric errors on regional femoral strain calculations was assessed using linear regressions and Root Mean Square (RMS) errors. Calculations were performed for each region along the length of the femur. The normality of the strain

difference distributions was assessed using Kolmogorov–Smirnov test (Lilliefors, 1967). The Student t-test (Hazewinkel, 1994) and Wilcoxon test (Wilcoxon, 1946) were used to compare normal and non-normal differences in strain distributions over the different activities.

The effect of sample size on inter-participant strain averages was assessed by calculating the regional average tensile and compressive strains using a sample size increasing from 2 to 10 participants. The linear regression and the RMS error between the inter-participant (sample size: 10) scaled-generic and image-based averages of regional femoral strains were also calculated.

3. Results

The joint angles, net joint moments, hip-joint contact forces, and muscle activation patterns calculated for walking using the image-based models were consistent with earlier findings (see Figs S1-S2 in Supplementary Material). The peak femoral strains in the proximal-lateral femoral shaft calculated for walking and stair ascent were consistent with corresponding strain measurements reported by Aamodt et al. (1997); mean peak tensile and compressive strains calculated for the ten participants ranged from 1351 to 1647 με and 971 to 988 με, respectively, compared to corresponding strains of 1198-1454 με and 393-948 με measured from two hip syndrome patients.

Scaled-generic and image-based anthropometric differences for the hip-to-hip and hip-to-knee distances were within ±1.04 cm (±6.1% of the hip-to-hip image-based distance) and ±1.88 cm (±5.5% of the hip-to-knee image-based distance), while the femoral anteversion and CCD angles were within ±8.9 and ±2.8 degrees, respectively, and femoral neck length was within ±0.4 cm (Table 2). The average absolute and percent differences in the moment arms of the hip- and knee-spanning muscles calculated for all six activities were -1.7 cm and -0.85% whereas the peak absolute and percent differences were 15.6 cm and +38.9% (Table

3). The linear regression between the scaled-generic and image-based muscle and hip contact forces yielded a coefficient of determination of R^2 = 0.78 for muscle forces and R^2 = 0.74-0.91 for the hip contact force components. The average Root Mean Squared Error (RMSE) ranged from 0.2-0.7 BW for the hip contact force components and was 0.1 BW for the muscle forces. The slope of the regression line ranged from 0.77-0.85 (0.76-0.86 95% confidence interval) for the hip contact force components and was 0.89 (0.88-0.89 95% confidence interval) for the muscle forces (Fig. 3) (see also Fig. S3). The median difference between scaled-generic and image-based moments was -8.6 Nm at the distal constraint and -1.1 Nm at the hip joint centre, while the 80th percentile of scaled-generic and image-based moment differences was -155.8 Nm at the distal constraint and -25.4 Nm at the hip centre (Fig. S4).

The coefficient of determination relating scaled-generic and image-based femoral strains decreased in the proximal-to-distal direction along the femur from level A to level H. The coefficient of determination varied from R^2 = 0.92 (level A, anterior) to R^2 = 0.48 (level H, medial). The average strain error (RMSE) varied from 380 με (level A, anterior) to 4,064 με (level H, medial). The peak strain error varied from 2,821 με (level A, anterior) to 34,166 με (level H, medial) (Fig. 4). The strain error distribution was not normally distributed (Kolmogorov–Smirnov, Lilliefors, $p<0.001$) and was activity-independent (Wilcoxon test, alpha = 0.05) (Fig. S5). Scaled-generic and image-based strain maps were different both in terms of the spatial distribution of strain and in magnitude. The differences in spatial distribution reached a peak at the most distal level H, at which point the location of the peak strain differed by as much as an anatomical quadrant compared to the image-based models (Fig. 5). The peak tensile and compressive strain differences per femoral level (A-H) increased linearly (R^2 = 0.77-0.82) from the proximal to distal femur, reaching 1051 με and -

570 με, respectively, in the femoral neck (levels A to C), and 12,307 με and -3,668 με in the remainder of the femur (levels D to H) (Fig. 6).

The inter-participant average for regional bone strain was a monotonic function of sample size that converged asymptotically (Fig. S6). The inter-participant averages for the scaled-generic and image-based bone strains showed similar patterns (Fig. 7); the coefficient of determination was $R^2 = 0.95$, the RMSE was 430 με, and the slope of the regression line was 0.96 (95% confidence interval: 0.96-0.97) (Fig. S7).

4. Discussion

We examined the sensitivity of femoral strain calculations to the anthropometric errors committed while scaling a generic musculoskeletal model to an individual participant's anatomy. Our results indicate that anthropometric errors cause a region-dependent strain error, which may lead to unrealistic participant-specific strain calculations in every femoral sub-region. In accordance with the central limit theorem, however, averaging the calculated bone strains over a cohort of participants can reduce strain errors, making scaled-generic models a viable tool for studying average patterns of femoral strains within a cohort of participants.

The anthropometric errors caused a region-dependent participant-specific strain error that increased from 2,821-5,500 με in the very proximal neck to 22,620-34,166 με in the distal diaphysis (Fig. 4). These region-dependent strain differences are attributable to scaled-generic and image-based differences in terms of hip contact force (Fig. 3), muscle forces (Fig. S3) and moments exerted on the femur by scaled-generic and image-based force systems (Fig. S4). Calculated strain values ranged from 39% to 468% of the bone yield strain threshold (i.e. 7,300 με in tension and 10,400 με in compression) reported by Bayraktar et al. (2004). Therefore, anthropometric errors in scaled-generic models may lead to unrealistic

estimates of participant-specific regional femoral strains. Specifically, image-based and scaled-generic strain maps over level-by-level femoral cross-sections differed either in terms of orientation or magnitude: orientation differences could cause the peak strain location to rotate about the femoral axis by up to a quadrant (Fig. 6), whereas peak strain differences over level-by-level cross-sections in the femoral neck (levels A to C) were -570 με in compression and 1051 με in tension (Fig. 5), overall less than the 14.4% of the yield strain reported by Bayraktar et al. (2004). Therefore, scaled-generic models may be used to calculate the participant-specific peak strain in the femoral neck when the peak strain, but not its location, is of interest.

The comparison of inter-participant averages of image-based and scaled-generic regional femoral strains showed good agreement for every femoral sub-region (Fig. 7). The average error was 430 με and the coefficient of determination was R^2=0.95. Therefore, scaled-generic models are a viable tool for determining average femoral strains within a cohort of participants. The minimum size of the cohort is a function of the femoral region of interest and the admissible error for the intended application, and can be determined using convergence plots (Figure S6).

The reliability of the present results can be better understood by comparing intermediate results with previous findings. Image-based models yielded joint kinematics, net joint moments, hip joint forces, muscle activation patterns and bone strains in the proximal-lateral femoral shaft in agreement with earlier studies (Aamodt et al., 1997; Bergmann et al., 2001; Inman et al., 1989; Kadaba et al., 1989; Stacoff et al., 2005). We found errors in the hip-joint-centre location of up to 2.01 cm for the scaled-generic model, which is similar to the 2.09 proximal shift of the hip-joint-centre location reported by Lenaerts et al. (2009). Errors in the flexion-extension moment arms of the hip-spanning muscles over the investigated activities were as high as 38.9% (Table 3), which agrees with the 36.3% error reported by Scheys et al.

(2008) for gait. The 0.52 BW difference between scaled-generic and image-based hip joint forces reported by Lenaerts et al. (2009) for walking compares well with the 0.2-0.7 BW average difference over a broader range of tasks found in the present study. Image-based models yielded a tensile strain of 1,912 με (Fig. 7) in the femoral neck during walking, in line with the 2,004 με reported earlier using a model entirely generated from dissection data (Martelli et al., 2014b).

There are limitations associated with the analyses presented. The imaging protocol was designed to focus only on the femur and pelvis to minimize the X-ray radiation dose given to participants. Extending the image-based anthropometric information to the remaining body segments may have increased further scaled-generic and image-based femoral strain differences. The reported average strain values might not be representative for larger cohorts due to the high strain errors (Fig. 4) and the limited sample size of 10 participants. Additional sources of error that can affect femoral strain calculations include the definition of the constraint of the femur (Cleather and Bull, 2011; Martelli et al., 2015), muscle function (Valente et al., 2012; Xiao and Higginson, 2010) and its changes while aging (Thelen, 2003). Functional methods have been found to improve the estimation of the hip joint centre (Leardini et al., 1999) and of the knee rotation axis (Schache et al., 2006) over landmark-based scaling procedures and have been used to determine musculoskeletal forces at the knee (Trepczynski et al., 2012). Therefore, functional methods may help reduce anthropometric errors in scaled-generic models and their effect on femoral strain calculation. Regarding the effect of aging on muscle function, Thelen (2003) concluded that age-related changes in muscle function may be important when simulating movements with substantial power requirement while Lim et al. (2012) showed that muscle function is invariant to age when walking speed is controlled. Therefore, we do not expect femoral strains during daily activities to be significantly affected by age-related changes in muscle function. Last, the

absence of in vivo bone deformation measurements makes it impossible to assess the accuracy of scaled-generic and image-based models. However, the present results provide information about the sensitivity of model outputs to anthropometric errors in scaled-generic musculoskeletal models.

Despite the above limitations, this study provides a better understanding of the sensitivity of femoral strain calculations to anthropometric errors committed while scaling a reference model to a participant's anatomy. Our analyses showed that the calculation of participant-specific bone strain from scaled-generic models should be considered with caution because it may yield unrealistic strain estimates, particularly in the most distal region. In accordance with the central limit theorem, however, the effect of anthropometric errors is reduced significantly by averaging strain calculations over multiple participants, making the use of scaled-generic models a viable solution with which to assess cohort-based averages of femoral strain during different activities.

5. Acknowledgments

This study was supported by an Australian Research Council Discovery Projects grant (DP1095366) and an Innovation Fellowship from the Victorian Endowment for Science, Knowledge, and Innovation to MGP. Funding provided by the Australian Research Council (DE140101530) to SM is also gratefully acknowledged.

6. References

Aamodt, A., Lund-Larsen, J., Eine, J., Andersen, E., Benum, P., Husby, O.S., 1997. In vivo measurements show tensile axial strain in the proximal lateral aspect of the human femur. Journal of Orthopaedic Research 15, 927–31.

Abdel Fatah, E.E., Shirley, N.R., Mahfouz, M.R., Auerbach, B.M., 2012. A three-dimensional analysis of bilateral directional asymmetry in the human clavicle. American Journal of Physical Anthropology 149, 547–59.

Ackland, D.C., Lin, Y.-C., Pandy, M.G., 2012. Sensitivity of model predictions of muscle function to changes in moment arms and muscle-tendon properties: a Monte-Carlo analysis. Journal of Biomechanics 45, 1463–71.

Anderson, F.C., Pandy, M.G., 2001. Static and dynamic optimization solutions for gait are practically equivalent Journal of Biomechanics 34, 153–161.

Bayraktar, H.H., Morgan, E.F., Niebur, G.L., Morris, G.E., Wong, E.K., Keaveny, T.M., 2004. Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. Journal of Biomechanics 37, 27–35.

Bergmann, G., Deuretzbacher, G., Heller, M., Graichen, F., Rohlmann, A., Strauss, J., Duda, G.N., 2001. Hip contact forces and gait patterns from routine activities. Journal of Biomechanics 34, 859–871.

Blemker, S.S., Asakawa, D.S., Gold, G.E., Delp, S.L., 2007. Image-based musculoskeletal modeling: applications, advances, and future opportunities. Journal of Magnetic Resonance Imaging 25, 441–451.

Cleather, D.J., Bull, a. M.J., 2011. Knee and hip joint forces - sensitivity to the degrees of freedom classification at the knee. Proceedings of the Institution of Mechanical Engineers, Part H. 225, 621–626.

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

Correa, T. A., Baker, R., Kerr Graham, H., Pandy, M.G., Graham, H.K., 2011. Accuracy of generic musculoskeletal models in predicting the functional roles of muscles in human gait. Journal of Biomechanics 44, 2096–2105.

Correa, T.A., Crossley, K.M., Kim, H.J., Pandy, M.G., 2010. Contributions of individual muscles to hip joint contact force in normal walking. Journal of Biomechanics 43, 1618–22.

Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G., 2007. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Transactions on Biomedical Engineering 54, 1940–1950.

Delp, S.L., Loan, J.P., Hoy, M.G., Zajac, F.E., Topp, E.L., Rosen, J.M., 1990. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. IEEE Transactions on Biomedical Engineering 37, 757–767.

Dorn, T.W., Schache, A.G., Pandy, M.G., 2012. Muscular strategy shift in human running: dependence of running speed on hip and ankle muscle performance. Journal of Experimental Biology 215, 1944–56.

Hazewinkel, M., 1994. Encyclopaedia of Mathematics (set), Encyclopaedia of Mathematics. Springer Netherlands. ISBN 1556080107

Inman, V.T., Ralston, H., Todd, F., 1981. Human Walking. MD: Williams & Wilkins, Baltimore. ISBN 068304348X

Jonkers, I., Lenaerts, G., Mulier, M., Van der Perre, G., Jaecques, S., 2008. Relation between subject-specific hip joint loading, stress distribution in the proximal femur and bone mineral density changes after total hip replacement. Journal of Biomechanics 41, 3405–3413.

Kadaba, M.P., Ramakrishnan, H.K., Wootten, M.E., Gainey, J., Gorton, G., Cochran, G. V, 1989. Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. Journal of Orthopaedic Research 7, 849–60.

Lang, T.F., Saeed, I.H., Streeper, T., Carballido-Gamio, J., Harnish, R.J., Frassetto, L.A., Lee, S.M.C., Sibonga, J.D., Keyak, J.H., Spiering, B.A., Grodsinsky, C.M., Bloomberg, J.J., Cavanagh, P.R., 2014. Spatial heterogeneity in the response of the proximal femur to two lower-body resistance exercise regimens. Journal of Bone and Mineral Research 29, 1337–45.

Leardini, A., Cappozzo, A., Catani, F., Toksvig-Larsen, S., Petitto, A., Sforza, V., Cassanelli, G., Giannini, S., 1999. Validation of a functional method for the estimation of hip joint centre location. Journal of Biomechanics. 32, 99–103.

Lenaerts, G., Bartels, W., Gelaude, F., Mulier, M., Spaepen, A., Van der Perre, G., Jonkers, I., 2009. Subject-specific hip geometry and hip joint centre location affects calculated contact forces at the hip during gait. Journal of Biomechanics 42, 1246–1251.

Lilliefors, H.W., 1967. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. Journal of the American Statistical Association 62, 399–402.

Lim, Y. P., Lin, Y.-C. & Pandy, M. G., 2013. Muscle function during gait is invariant to age when walking speed is controlled. Gait Posture 38, 253–9.

Martelli, S., Kersh, M.E., Schache, A.G., Pandy, M.G., 2014a. Strain energy in the femoral neck during exercise. Journal of Biomechanics 47: 1784-1791.

Martelli, S., Pivonka, P., Ebeling, P.R., 2014b. Femoral shaft strains during daily activities: implications for Atypical Femoral Fractures. Clinical Biomechanics (Bristol, Avon) 29, 869–876.

Martelli, S., Valente, G., Viceconti, M. & Taddei, F., 2015. Sensitivity of a subject-specific musculoskeletal model to the uncertainties on the joint axes location. Computer Methods in Biomechanics and Biomedical Engineering 18, 1555–63.

Morgan, E.F., Bayraktar, H.H., Keaveny, T.M., 2003. Trabecular bone modulus-density relationships depend on anatomic site. Journal of Biomechanics 36, 897–904.

Redl, C., Gfoehler, M., Pandy, M.G., 2007. Sensitivity of muscle force estimates to variations in muscle-tendon properties. Human Movement Science 26, 306–319.

Schache, A.G., Baker, R., Lamoreux, L.W., 2006. Defining the knee joint flexion-extension axis for purposes of quantitative gait analysis: an evaluation of methods. Gait Posture 24, 100–109.

Scheys, L., Loeckx, D., Spaepen, A., Suetens, P., Jonkers, I., 2009. Atlas-based non-rigid image registration to automatically define line-of-action muscle models: a validation study. Journal of Biomechanics 42, 565–72.

Scheys, L., Van Campenhout, A., Spaepen, A., Suetens, P., Jonkers, I., 2008. Personalized MR-based musculoskeletal models compared to rescaled generic models in the presence of increased femoral anteversion: effect on hip moment arm lengths. Gait Posture 28, 358–365.

Schileo, E., Taddei, F., Malandrino, A., Cristofolini, L., Viceconti, M., 2007. Subject-specific finite element models can accurately predict strain levels in long bones. Journal of Biomechanics 40, 2982–9.

Stacoff, A., Diezi, C., Luder, G., Stüssi, E., Kramers-de Quervain, I.A., 2005. Ground reaction forces on stairs: effects of stair inclination and age. Gait Posture 21, 24–38.

Taddei, F., Martelli, S., Reggiani, B., Cristofolini, L., Viceconti, M., 2006. Finite-element modeling of bones from CT data: sensitivity to geometry and material uncertainties. IEEE Transactions on Biomedical Engineering 53, 2194–200.

Taddei, F., Martelli, S., Valente, G., Leardini, A., Benedetti, M.G., Manfrini, M., Viceconti, M., 2012. Femoral loads during gait in a patient with massive skeletal reconstruction. Clinical Biomechanics (Bristol, Avon) 27, 273–280.

Thelen, D.G., 2003. Adjustment of muscle mechanics model parameters to simulate dynamic contractions in older adults. Journal of Biomechanical Engineering 125, 70–77.

Valente, G., Martelli, S., Taddei, F., Farinella, G., Viceconti, M., 2012. Muscle discretization affects the loading transferred to bones in lowerlimb musculoskeletal models. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 226, 161–9.

Van Rietbergen, B., Huiskes, R., Eckstein, F., Ruegsegger, P., 2003. Trabecular bone tissue strains in the healthy and osteoporotic human femur. Journal of Bone and Mineral Research 18, 1781–1788.

Viceconti, M., Olsen, S., Nolte, L.P.P., Burton, K., 2005. Extracting clinically relevant data from finite element simulations. Clinical Biomechanics 20, 451–454.

Wilcoxon, F., 1946. Individual comparisons of grouped data by ranking methods. Journal of Economical Entomology 39, 269.

461 Wu, G., Siegler, S., Allard, P., Kirtley, C., Leardini, A., Rosenbaum, D., Whittle, M.,
462 D'Lima, D.D., Cristofolini, L., Witte, H., Schmid, O., Stokes, I., 2002. ISB
463 recommendation on definitions of joint coordinate system of various joints for the
464 reporting of human joint motion--part I: ankle, hip, and spine. Journal of Biomechanics
465 35, 543–548.

466 Xiao, M., Higginson, J., 2010. Sensitivity of estimated muscle force in forward simulation of
467 normal walking. Journal of Applied Biomechanics 26, 142–149.

468

469

470

471 **Table 1** – Participant details (all female).

Participant	**Age (years)**	**Weight (kg)**	**Height (cm)**	**BMI (kg/m^2)**
1	74	51	150	22.7
2	64	52	150	23.1
3	72	66	158	26.6
4	68	61	158	24.6
5	68	53	159	21.0
6	60	85	153	36.3
7	60	96	170	33.1
8	64	69	168	24.6
9	64	71	165	26.1
10	73	59	157	23.9

472 BMI = Body Mass Index.

473

474 **Table 2** – Differences between scaled-generic and image-based hip-to-hip and hip-to-knee
475 distances, femoral anteversion angle, caput-collum-diaphyseal (CCD) angle and femoral neck
476 length.

Participant #	D hip-to-hip distance cm (%)	*Hip-to-hip distance mm	D hip-to-knee distance cm (%)	*Hip-to-knee distance mm	D Femoral anteversion deg (%)	*Femoral anteversion deg	D CCD angle deg (%)	*CCD angle deg	D neck length cm (%)	*Neck length cm
1	0.16 (1.0)	161	0.70 (2.1)	342	10.8 (65)	16.7	0.5 (0.4)	121.1	2.5 (4.9)	4.9
2	-1.29 (-8.1)	160	1.44 (4.4)	329	25.2 (1070)	2.3	-4.3 (-3.4)	125.9	7.5 (14.9)	4.3
3	-0.06 (-0.4)	180	-3.24 (-8.7)	372	6.8 (33)	20.7	-8.5 (-6.6)	130.1	-4.2 (-8.3)	5.5
4	0.50 (2.9)	170	-1.42 (-4.1)	343	12.9 (89)	14.6	-7.0 (-5.5)	128.6	-0.5 (-1.1)	5.0
5	-0.38 (-2.1)	180	-0.18 (-0.5)	368	21.9 (392)	5.6	-4.5 (-3.6)	126.1	3.0 (5.3)	5.3
6	2.01 (11.6)	174	-2.56 (-7.6)	339	20.1 (270)	7.4	-4.8 (-3.8)	126.4	-1.3 (-2.8)	4.8
7	1.38 (7.5)	183	1.57 (4.2)	375	21.9 (392)	5.6	-4.5 (-3.6)	126.1	3.4 (6.1)	5.3
8	-0.71 (-3.9)	180	1.14 (0.4)	364	20.3 (283)	7.2	-4.1 (-3.3)	125.7	0.0 (0.1)	5.4
9	-0.68 (-3.7)	183	0.99 (2.7)	369	1.8 (7)	25.7	-3.5 (-2.8)	125.1	-4.1 (-7.5)	5.9
10	-0.90 (-5.5)	165	2.58 (8.4)	306	-1.9 (-6)	29.4	4.3 (3.7)	117.3	-3.8 (-7.7)	5.3
Mean	0.001 (-0.07)	174	0.001 (0.12)	351	14.0 (259)	13.5	-3.7 (-2.8)	125.3	0.2 (0.4)	5.2
SD	1.04 (6.1)	9	1.88 (5.5)	23	8.9 (308)	8.9	3.5 (2.8)	3.5	3.7 (7.1)	0.4

477 Percentage differences between image-based and scaled-generic lengths are expressed as a percentage of the
478 corresponding image-based length. Reported are the mean values with standard deviations given in parentheses.
479 * The image-based parameters used as reference.

480

481

Table 3 - Differences between scaled-generic and image-based moment arms of hip- and knee-spanning muscles calculated over the six studied activities.

Muscle name	Degree of freedom	Average difference in mean moment arm (± SD) (%)	Mean image-based moment arm (mm)	95% limits of agreement (Bland-Altman) (mm)
Biceps femoris long head	Hip flexion	7.7 (18.2)	-60	-29:11
Gluteus maximus anterior		6.9 (12.8)	-63	-25:7
Gluteus maximus middle		3.1 (10.4)	-66	-16:10
Gluteus maximus posterior		6.2 (10.4)	-74	-16:11
Iliacus		37.6 (13.4)	41	7:25
Psoas major		38.9 (15.8)	40	7:29
Rectus femoris		11.7 (8.9)	43	-2:13
Semimembranosus		6.0 (21.4)	-53	-34:12
Semitendinosus		11.2 (17.2)	-60	-37:7
Adductor brevis	Hip adduction	-8.5 (5.4)	72	-12:2
Adductors longus		-6.8 (6.1)	74	-13:5
Adductor magnus prox.		-0.9 (6.0)	78	-12:8
Adductor magnus middle		7.3 (9.3)	67	-7:15
Adductor magnus distal		8.3 (19.7)	33	-4:21
Gluteus medius anterior		9.2 (17.9)	-45	-19:8
Gluteus medius middle		3.1 (10.2)	-44	-12:5
Gluteus medius posterior		-1.1 (12.3)	-36	-7:6
Gluteus minimus anterior		17.8 (18.4)	-37	-21:1
Gluteus minimus middle		12.2 (11.4)	-39	-13:1
Gluteus minims posterior		12.3 (8.3)	-35	-11:0
Gracilis		-0.3 (7.4)	59	-5:10
Tensor fascia latae		8.6 (17.1)	-46	-16:15
Gemelli	Hip rotation	11.7 (9.4)	-31	-11:2
Pectineus		-109.8 (61.4)	-4	1:9
Perineus		7.8 (19.3)	-29	-13:5
Quadratus femoris		15.0 (18.6)	-37	-20:4
Biceps femoris long head	Knee extension	-5.4 (16.4)	-31	-15:11
Biceps femoris short head		-3.4 (16.6)	-30	-14:10
Lateral gastrocnemius		12.7 (27.9)	-25	-29:6
Medial gastrocnemius		-0.4 (21.3)	-27	-11:10
Rectus femoris		-20.2 (11.1)	52	-27:4
Semimembranosus		-3.7 (13.9)	-37	-16:10
Semitendinosus		-1.7 (12.6)	-43	-16:10
Vastus intermedius		-20.6 (11.1)	52	-27:3
Vastus lateralis		-21.4 (15.0)	52	-32:1
Vastus medialis		-20.5 (12.4)	52	-29:5
Mean		0.85 (15.1)	-6.6	

Muscle moment arm differences were calculated as scaled-generic minus image-based values over the six investigated activities, averaged for both limbs and expressed as a percentage of the mean value of the image-based muscle moment arm, which is also reported. SD = standard deviation.

488
489 **Figure Captions**

490 Fig. 1 – Generic model (A), Scaled-generic model (B), and image-based model (C) used
491 in this study. Pink spheres (panels A, B, and C) represent virtual markers attached to the
492 model. Blue markers (panel B) are the skin-mounted markers used in the gait experiments.
493 The markers encircled in red (panel B) were used to calculate the characteristic segment
494 lengths used to scale the generic model. The distances indicated are as follows: (1) sacrum to
495 seventh cervical spine; (2) acromium to elbow; (3) elbow to wrist; (4) the span of anterior
496 superior iliac spine; (5) anterior superior spine to lateral epicondyle; (6) lateral epicondyle to
497 lateral malleolus; (7) heel to toe. The inset to the model in panel C shows the solid models of
498 the femur and pelvis segments created from the CT images obtained from each participant.
499 The CT images were used to identify the knee, the hip and the sacrum joints (red marker) and
500 the muscle paths depicted in blue.

501 Fig. 2 – Finite-element model of the femur (right) created from the CT images obtained of
502 the femur and pelvis segments (left). Femoral strains were analysed at 8 different levels along
503 the length of the femur. Each level was sub-divided into 4 aspects (anterior, posterior, medial
504 and lateral) resulting in 32 sub-regions. The colour scale represents the distribution of the
505 values of Young's modulus as calculated from the CT images.

506 Fig. 3 – Linear regression analysis between the scaled-generic and image-based models
507 for the hip joint force components and magnitude (R^2 = correlation coefficient, b = slope of
508 regression line, CI = confidence intervals associated with the slope, RMSE = root mean
509 squared error, MAX ERROR = maximum error between the scaled-generic and image-based
510 models).

511 Fig. 4 – Linear regressions between the scaled-generic and image-based strains over the
512 32 femoral sub-regions. R^2 = correlation coefficient, b = slope of regression line, CI =

513 confidence intervals associated with the slope, RMSE = root mean square error, MAX
514 ERROR = maximum error between the scaled-generic and image-based models.

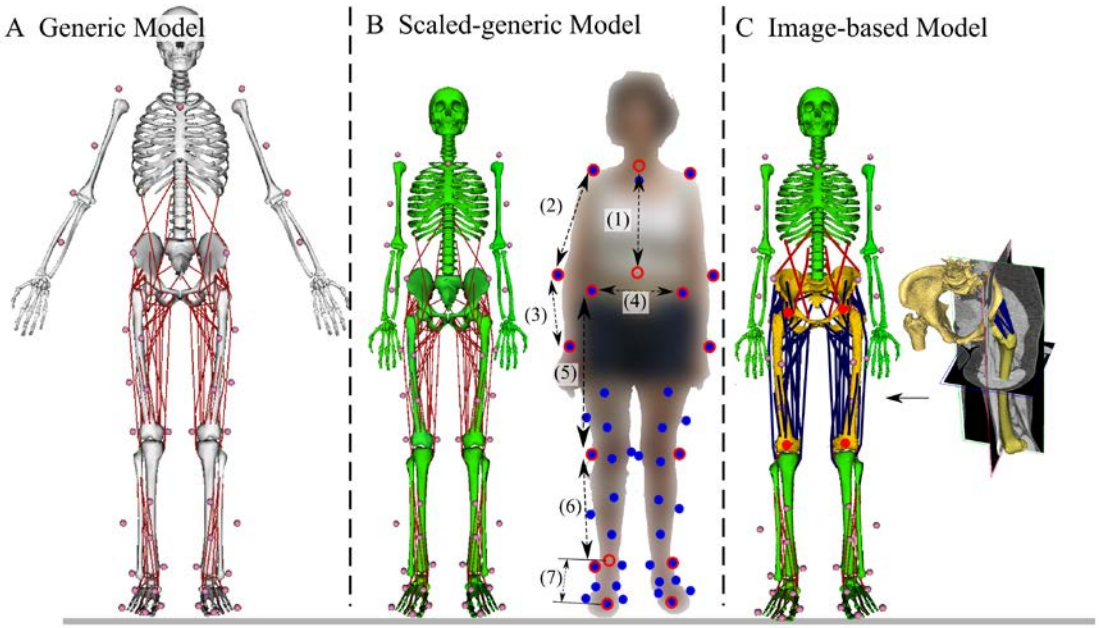
515 Fig. 5 –The distribution of principal tensile strains in the scaled-generic (left) and image-
516 based (right) are shown at the most distal femoral level considered (level H) for a single
517 participant during the late stance phase of walking. The peak strain in the scaled-generic
518 model is in the posterior aspect of the femur, while the peak strain in the image-based model
519 is seen on the lateral aspect.

520 Fig. 6 – Linear regression analysis for the errors in the peak tensile (top) and compressive
521 (bottom) strains shown at the different levels of the femur (levels A-H). Each data point (blue
522 diamond) represents the average of the peak strain error while the error bar represents the
523 95% limits of agreement (Bland-Altman). R^2 = correlation coefficient, x = femoral level, y =
524 peak tensile or compressive strain error.

525 Fig. 7 – Regional inter-participant strains (i.e., cohort average, principal tensile (red) and
526 compressive (blue) strains) calculated from both the scaled-generic (dashed lines) and image-
527 based (solid lines) models for the stance phase of walking.

528

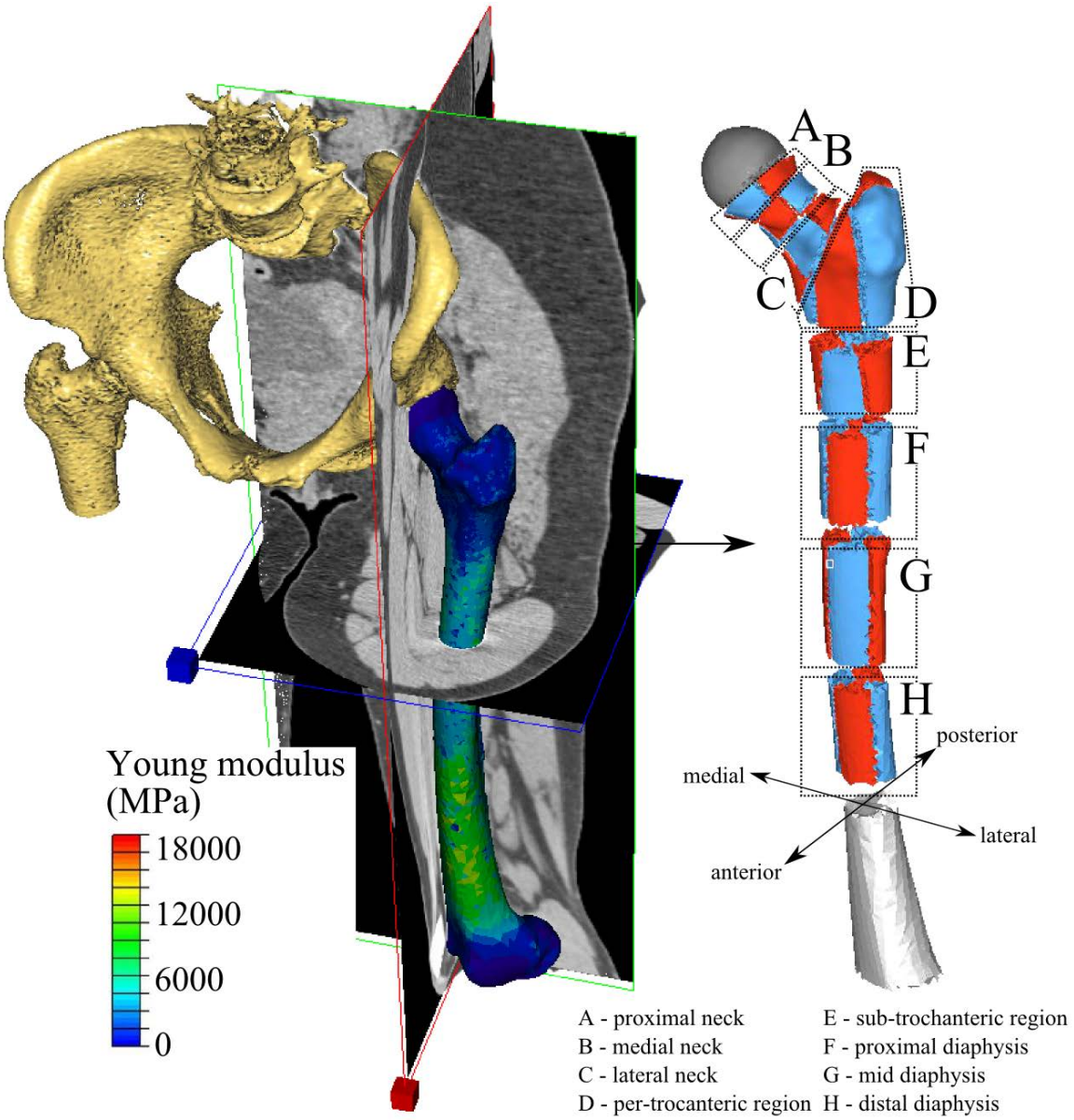
529 Figure 1



530

531

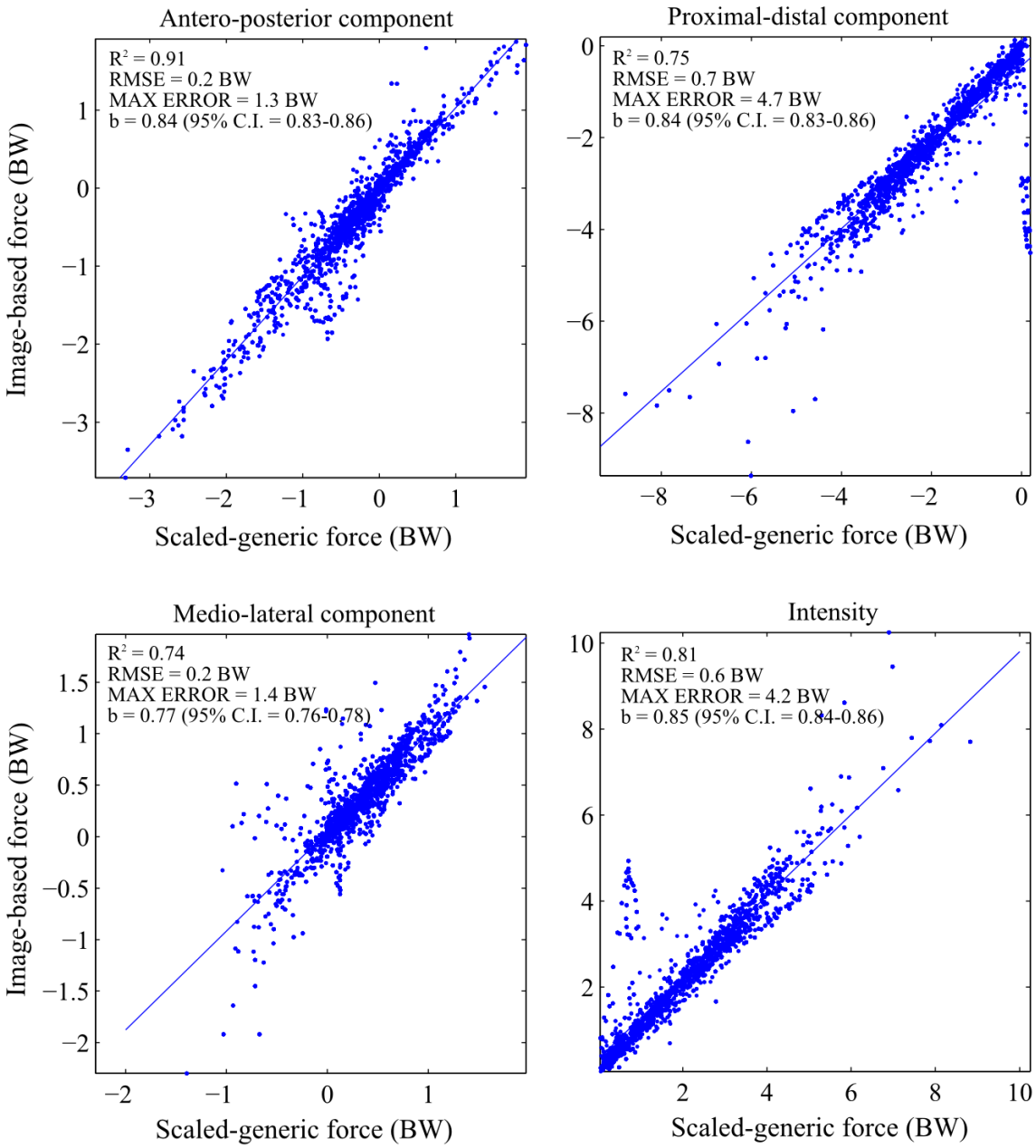
532 Figure 2



533

534

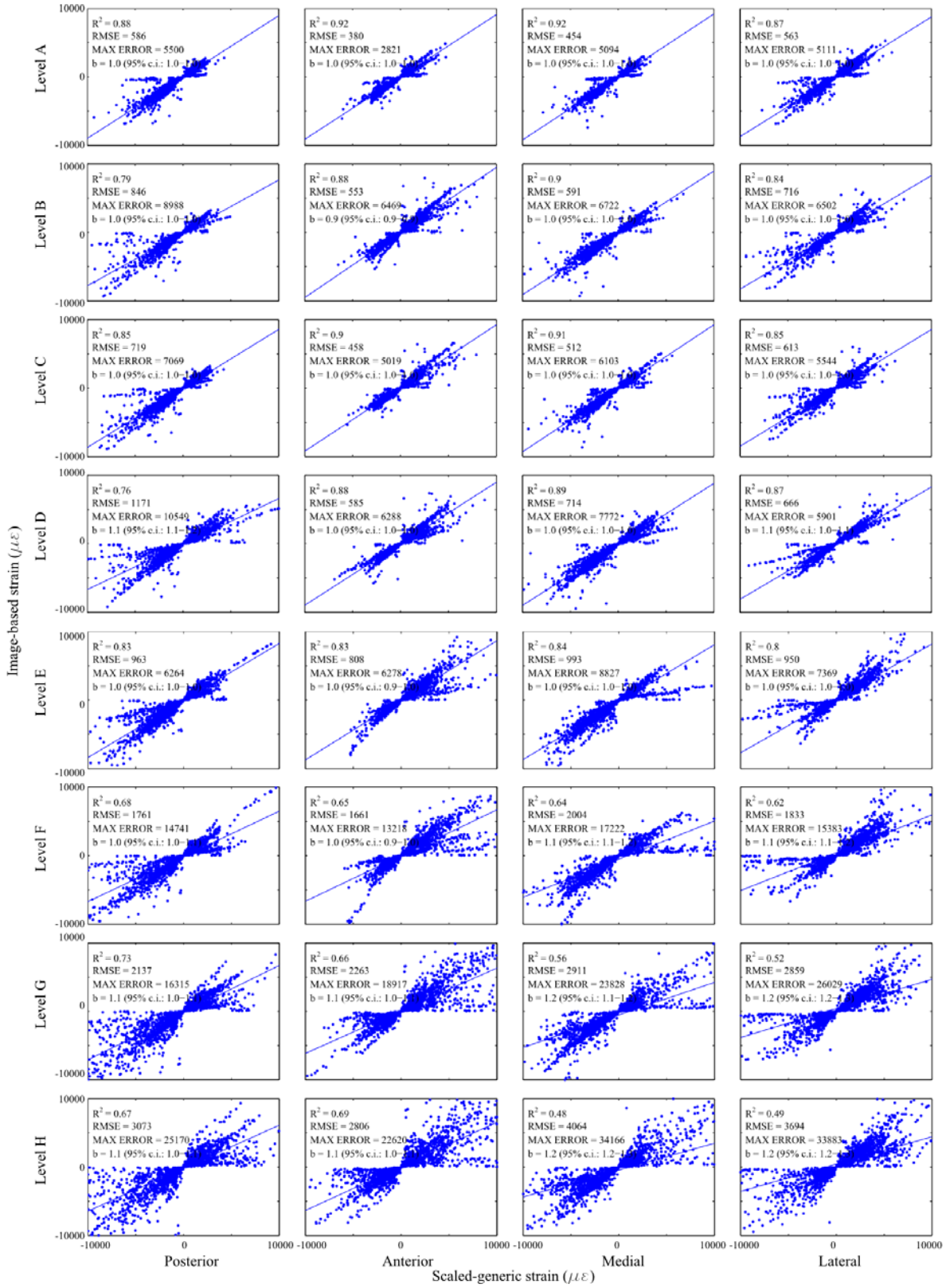
535 Figure 3



536

537

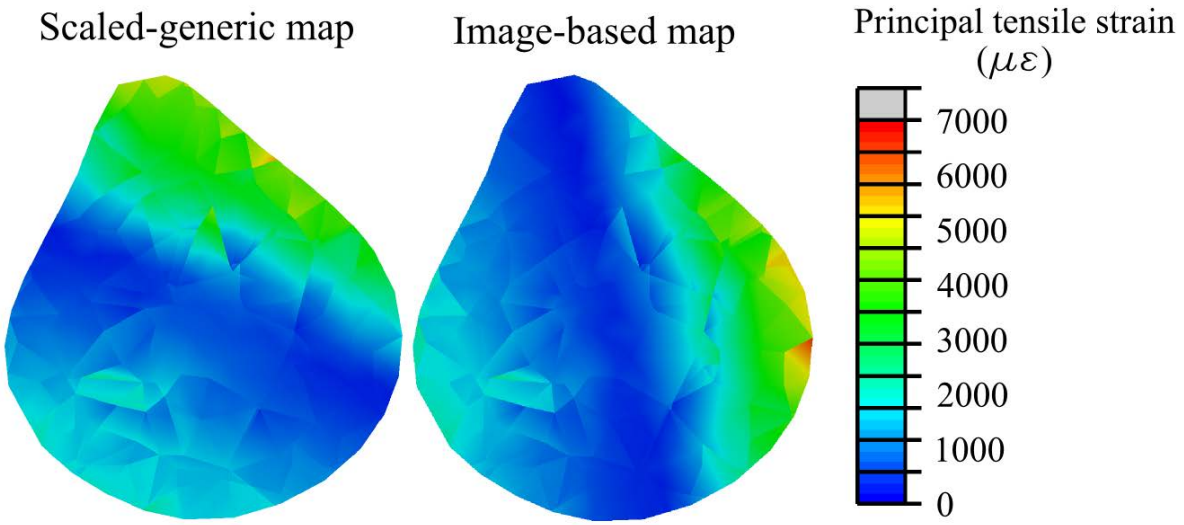
538 Figure 4



539

540

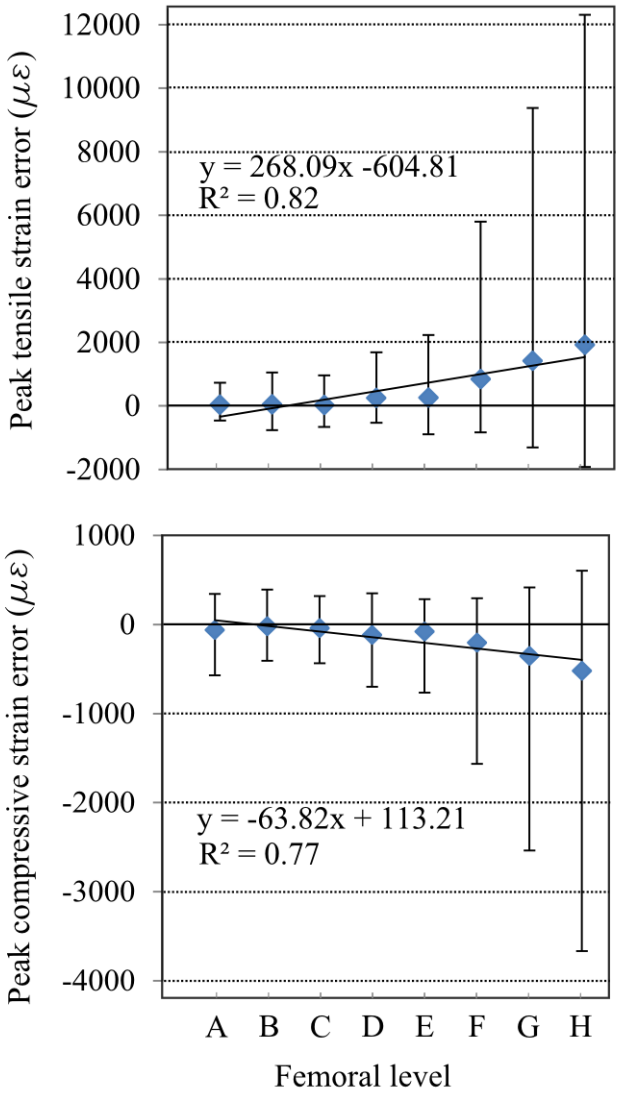
541 Figure 5



542

543

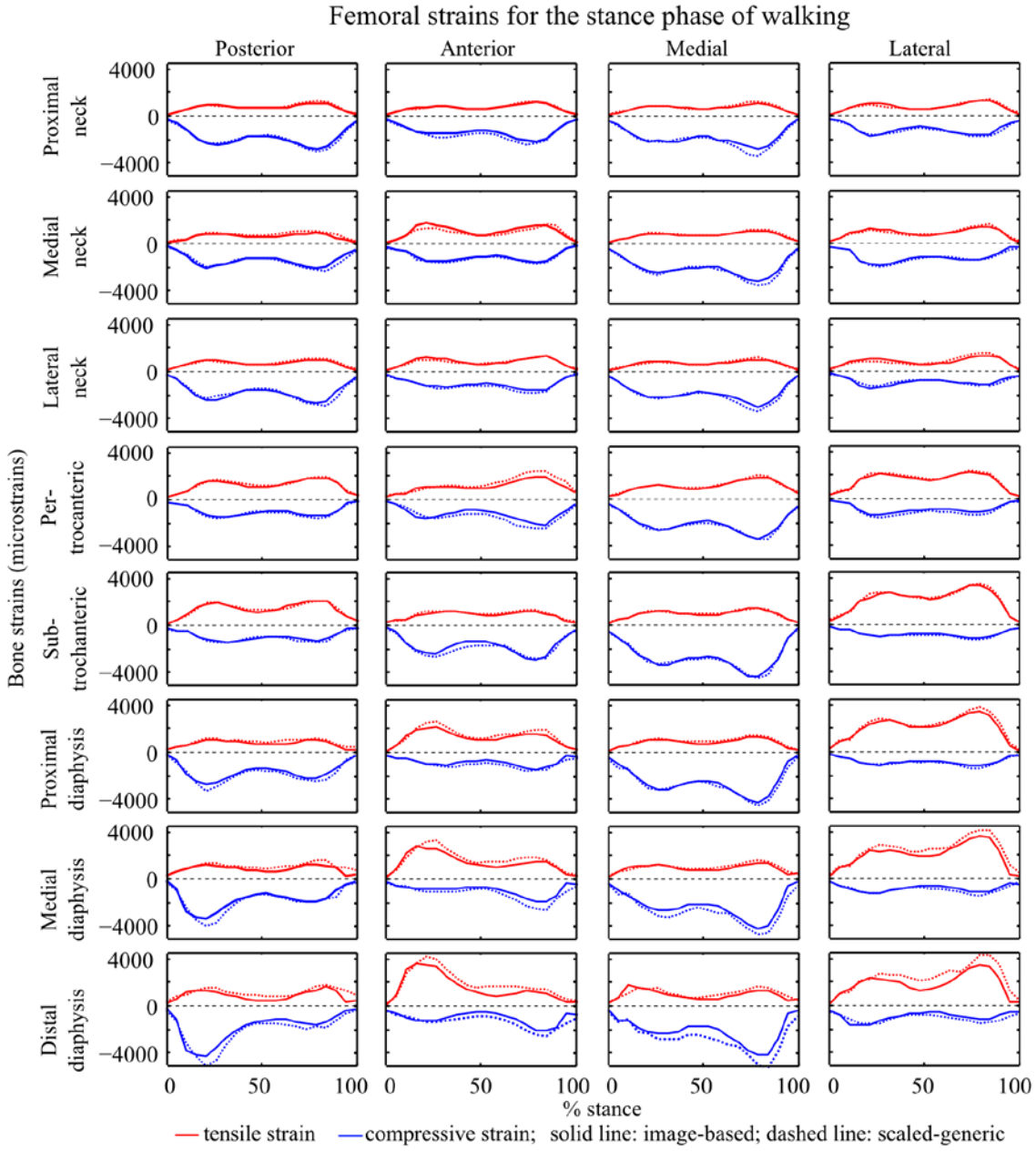
544 Figure 6



545

546

547 Figure 7



548