A statistical shape model of the tibia-fibula complex: sexual dimorphism and effects of age on reconstruction accuracy from anatomical landmarks

Olivia L. Bruce[a, b, c*], Michael Baggaley[a, c], Lauren Welte[d], Michael J. Rainbow[d], and W. Brent Edwards[a, b, c]

[a]*Human Performance Laboratory, Faculty of Kinesiology, University of Calgary, Calgary, Canada;* [b]*Biomedical Engineering Graduate Program, University of Calgary, Calgary, Canada;* [c]*McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, Canada* [d]*Department of Mechanical and Materials Engineering, Queen's University, Kingston, Canada*

*Corresponding author: Olivia L. Bruce, Human Performance Laboratory, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada, T2N 1N4. email: olivia.bruce@ucalgary.ca

This is an Accepted Manuscript of an article published by Taylor & Francis in Computer Methods in Biomechanics and Biomedical Engineering on 3 November 2021, available online: https://www.tandfonline.com/doi/full/10.1080/10255842.2021.1985111

A statistical shape model of the tibia-fibula complex: sexual dimorphism and effects of age on reconstruction accuracy from anatomical landmarks

A statistical shape model was created for a young adult population and used to predict tibia and fibula geometries from bony landmarks. Reconstruction errors with respect to CT data were quantified and compared to isometric scaling. Shape differences existed between sexes. The statistical shape model estimated tibia-fibula geometries from landmarks with high accuracy (RMSE = 1.51-1.62 mm), improving upon isometric scaling (RMSE = 1.78 mm). Reconstruction errors increased when the model was applied to older adults (RMSE = 2.11-2.17 mm). Improvements in geometric accuracy with shape model reconstruction changed hamstring moment arms 25-35% (1.0-1.3 mm) in young adults.

Keywords: lower extremity; bone; musculoskeletal model; participant-specific

Introduction

Musculoskeletal models are commonly used to estimate muscle forces and joint kinematic and kinetic parameters associated with human movement. Outputs from musculoskeletal models (e.g., muscle moment arms and joint contact forces) are highly sensitive to bone geometry (Scheys et al. 2008; Gerus et al. 2013; Clouthier et al. 2019; Ding et al. 2019; Xu et al. 2020), which is frequently defined using either participant-specific imaging or model scaling approaches. Advanced imaging, including computed tomography (CT) and magnetic resonance imaging, is the gold standard for quantifying participant-specific bone geometry. Of course, CT imaging requires ionizing radiation and both imaging modalities are costly and can be challenging to acquire. Consequently, it is more common to scale a 'generic' musculoskeletal model according to gross anthropometric measurements (Delp et al. 1990; Arnold et al. 2010), which does not necessarily capture potentially important differences in bone geometry among individuals.

Statistical shape models (SSM) provide an alternative method to incorporate participant-specific bone geometry into musculoskeletal models. A SSM numerically calculates the average and principal modes of variation of a shape (e.g., bone geometry) from a training set of models (Audenaert et al. 2019). These models can be used to reconstruct participant-specific geometry from incomplete information by morphing the average shape along the modes of variation to best fit the target data. In this way, the pelvis, femur, tibia-fibula complex, and multiple bones of the foot have been reconstructed from select anatomical landmarks identified either on the bone surface or using skin-mounted motion capture markers (Zhang et al. 2016; Grant et al. 2020; Nolte et al. 2020).

Bone geometry is known to vary as a function of age and sex, among other factors (Ruff & Hayes 1988; Stevens & Vidarsdóttir 2008; Mahfouz et al. 2012; Li et al. 2014; Brzobohatá et al. 2015; Brzobohatá et al. 2016; Audenaert et al. 2019). Indeed, periosteal expansion increases with age, although to a lesser degree in females than males (Ruff & Hayes 1988; Jee 2001). At the tibia, observed geometric differences between males and females include: greater protrusion of the tibial tuberosity, diaphyseal curvature, diaphyseal cross-sectional properties (Smock et al. 2009; Feldman et al. 2012; Sherk et al. 2012), condyle size (Mahfouz et al. 2012; Audenaert et al. 2019; Tümer et al. 2019), and metaphyseal slope (Brzobohatá et al. 2015; Brzobohatá et al. 2016). These differences have been used to classify bones into age and sex- specific groups with 61.0-98% accuracy, depending on the study (Stevens & Vidarsdóttir 2008; Brzobohatá et al. 2015; Brzobohatá et al. 2016; Audenaert et al. 2019).

Previous studies reconstructing the tibia-fibula complex from anatomical markers used participant groups with wide age ranges (20-70 years in Nolte et al. (2020) and 15-92 years in Zhang et al. (2016)) to train and test the SSMs. While these SSMs may be more widely generalizable, a model more specific to a population of interest such as young healthy adults - a commonly used group in biomechanics modelling studies - may provide more accurate results. Furthermore, these studies have not investigated how errors in shape influence musculoskeletal modelling parameters. Thus, the primary purpose of this study was to evaluate the accuracy of tibia-fibula complex reconstructions from anatomical bony landmarks using a SSM developed for a healthy young adult population, and quantify the subsequent effects on muscle moment arms. Errors associated with isometrically scaling the average model were also quantified for comparison. In line with previous literature (Zhang et al. 2016; Nolte et al. 2020), we hypothesized that reconstruction errors would be smaller for the SSM-generated geometries compared to isometric scaling. The secondary objectives of this study were (1) to quantify sex differences in bone geometry, and (2) to evaluate the generalizability of the developed young adult SSM to older adults. We expected to observe differences in size and shape between young adult males and females. Due to age-related differences in bone geometry, we hypothesized that reconstruction errors and changes in muscle moment arms would be larger for the older adults.

Materials and methods

Model development:

Forty-one physically active participants (22 F and 19 M, 18-23 years, physically active at least three times per week) were recruited to obtain a range of statures (mean (range), female: 1.66 m (1.49 – 1.80 m), 59.7 kg (47.7-71.8 kg), male: 1.77 m (1.62 – 1.87 m), 71.8 kg (60.0 – 83.7 kg)). CT scans of the left lower leg were obtained using a GE Revolution GSI (GE Healthcare, Waukesha, WI) with image acquisition settings of 120 kVp and 180 mA. Images were reconstructed with an in-plane resolution of 0.488 mm x

0.488 mm and a slice thickness of 0.625 mm. Ethics approval was obtained from the university's Conjoint Health Research Ethics Board and written, informed consent was obtained from each participant prior to scanning.

The tibia and fibula geometries were segmented using a semi-automatic procedure and surface meshes were created in the Mimics Innovation Suite (v21, Materialise, Leuven, Belgium). Nodal correspondence and registration were performed in MATLAB (R2020a, Mathworks, MA, USA). A template mesh was selected, corresponding to an individual with tibia/fibula surface area close to the sample mean, and contained 3874 and 2111 nodes for the tibia and fibula, respectively. A sensitivity analysis, evaluating shape errors resulting from the template deformation step, was used to determine the number of nodes for the tibia and fibula. Nodal correspondence between meshes was established using the Coherent Point Drift algorithm (Myronenko & Song 2010). This algorithm performs translation, rotation, scaling, and local deformation to match a moving point-set (template) to a fixed point-set (participant surface). A nearest neighbours algorithm was used to identify corresponding points. A preliminary analysis of nodal correspondence registration errors and the number of principal components needed to explain 95% of the variance demonstrated these measures were insensitive to the choice of template. Tibia and fibula point clouds were then combined and rigidly aligned using a generalized Procrustes analysis that retained bone size. The tibia and fibula were modelled together to include relative positioning between the two bones within the model.

A principal component analysis (PCA) was applied to the registered data to obtain the average shape and modes of variation (i.e., principal components) for the sample. An analysis described by Mei *et al*. (2008) evaluating bootstrap stability on mode direction and comparison with noise was used to determine the number of principal components to retain. Eight principal components accounting for 96.2% of the total variance in the model were ultimately retained. Scores for each retained principal component were compared using unpaired t-tests to determine if and how size and shape differed between sexes (SPSS v.26, IBM, NY, USA, $\alpha = 0.05$). Centroid size, the square root of the sum of squared Euclidean distances of all points in a shape from the centroid of the shape, was calculated. Pearson correlations were used to evaluate whether principal component scores were correlated with size.

Landmark-based reconstruction:

The tibia-fibula SSM meshes were reconstructed based on two sets of anatomical landmarks (Figure 1). The first set contained nine landmarks that could be identified through palpation, and thus used to estimate tibia and fibula geometries from skin-mounted motion capture markers: tibial tuberosity, medial condyle, lateral and medial malleoli, lateral aspect of the head of the fibula, anterior border of the tibia at 25%, 50% and 75% of the distance between the medial condyle and malleolus markers, lateral fibula diaphysis at 25% of the distance from the lateral malleolus to the lateral point on the head of the fibula. The second set contained the nine "palpable" landmarks described above as well as another five landmarks that could not be palpated, but could be captured through imaging, to determine if additional information regarding dimension and curvature improved reconstruction accuracy: posterior aspect of the medial condyle, posterior aspect of the mid-diaphysis (50%) of the tibia, apex of the fibular head, fibular diaphysis at 50% (posterior) and 85% (anterior) of the distance between the lateral malleolus and the lateral point on the fibular head. Both sets of landmarks (Figure 1) were manually digitised on the average shape as well as each participants' CT-based bone surface meshes in MATLAB (R2020a, Mathworks, MA, USA).

A leave-one-out analysis was performed, where each participant was removed from the SSM and reconstructed from the digitised landmarks. The average point cloud was fit to the participant's landmarks using rigid-body rotation, translation, and deformation along the principal components using a Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm (Liu & Nocedal 1989). The objective of the optimization was to minimize the squared Euclidean distance between reconstructed and target landmarks, which was performed using each set of landmarks (Fig 1). Tibia/fibula geometries were also generated by isometrically scaling the average point cloud. The scaling factor was the ratio of the distance between the lateral malleolus and lateral fibular head markers for the average shape and target data.

Generalizability to older individuals:

A dataset of bilateral lower-limb CT scans from 10 human cadavers (81.8 ± 10.7 years, 6 male, 4 female) was used to examine the robustness of the SSM when applied to an entirely new sample. Scans were performed using a GE Revolution GSI (GE Healthcare), with image acquisition settings of 120 kVp, 103 mA, in-plane resolution of 0.67 mm x 0.67 mm, and slice thickness of 0.625 mm. The same procedures described above were used to segment and generate surface meshes of the left tibia and fibula. In one of the scans, part of the left tibia/fibula was outside of the field-of-view. In this case, the right tibia/fibula geometries were segmented and mirrored. Landmarks, previously defined (Figure 1), were digitised on the bone surfaces. Tibia and fibula geometries were reconstructed using nine and fourteen landmarks with the optimization procedure described above. The average shape from the shape model was also isometrically scaled.

Musculoskeletal modelling

A musculoskeletal model was implemented in MATLAB (Mathworks, Natick, MA) and used to obtain moment arms of eighteen muscles attaching to the tibia and fibula (Supplementary Table S3). Initial geometry and muscle parameters were obtained from Arnold et al. (2010). The average shape from the SSM was rigidly aligned and scaled to the generic tibia and fibula of the musculoskeletal model. The surface nodes corresponding to muscle origin or insertion points were determined using a nearest neighbours algorithm.

The musculoskeletal model was scaled using markers that were digitized on the model at the lateral malleolus and head of the fibula. Reconstructed and CT-based surfaces were rigidly aligned to the scaled musculoskeletal model. The musculoskeletal model was then moved through a physiologic range of motion about the flexion-extension axis at the knee (0° to 142° flexion) and the ankle (15° dorsiflexion to -62° plantarflexion) (Soucie et al. 2011). Translations at the knee along the anterior-posterior and longitudinal axes changed as a function of knee flexion; no translations were allowed at the ankle. Muscle moment arms were computed using the tendon excursion method. The model was positioned at each joint angle, and then perturbed by ± 10°. The moment arm of each muscle was calculated from the change in muscle length divided by the change in joint angle. Patellar ligament moment arm was calculated as the perpendicular distance from the knee joint centre to the line of action of the ligament. Maximum difference in moment arm compared to the model using the CT-based geometry was calculated.

Statistics:

Reconstructed geometries were aligned with the participants' CT-based surface meshes using a rigid iterative closest points algorithm. A nearest-neighbour algorithm was used for each node to calculate RMSE and maximum distance error. Jaccard index, a measure of volumetric similarity - where values range from 0 (no similarity) to 1 (identical) - was also computed (Real & Vargas 1996). RMSE and maximum error were also calculated for the tibia and fibula separately, and for proximal, diaphysis, and distal regions (0-20%, 20-80%, and 80-100% of the axial length, respectively (Edwards et al. 2013)). Statistical tests were performed using SPSS (v26, IBM, NY, USA). In the leave-one-out analysis, error measures for at least one of the reconstruction methods did not meet the assumption of normality as defined by the Shapiro-Wilk test. Therefore, related-samples Friedman's analysis of variance tests (ANOVA) were used to evaluate differences in error measurements between reconstruction methods (9 landmarks, 14 landmarks, isometric scaling); pairwise comparisons were used when appropriate. For the older adult dataset, all error measures met the assumption of normality. In this case, repeated measures ANOVAs were used to evaluate differences in error measurements between reconstruction methods; again, pairwise comparisons were used when appropriate. Critical values for statistical tests were adjusted for multiple comparisons using Bonferroni corrections to maintain a familywise error rate of $\alpha = 0.05$.

Results

The first principal component in the SSM primarily captured differences in overall size and explained 79.8% of the total variance in the model (Table 1). The first eight components explained 96.2% of the variance (Table 1).

Sexual dimorphism:

Principal component 1 discriminated between males and females ($t = 4.727$, $p < 0.001$, whereby the tibia and fibula were larger in males. Principal component 1 was the only mode correlated with centroid size, with $r^2 = 0.99$. Sex differences in principal components 4 and 7 trended toward significance ($t = -2.231$, and -2.279, $p = 0.031$, and 0.029, respectively; Bonferroni-adjusted critical p-value = 0.05/8 comparisons = 0.006;

Figure 2). These principal components described shape differences including larger proximal epiphyseal regions, a more prominent tibial tuberosity, and more acute anterior-posterior curvature in the tibia in males compared to females (Figure 2).

Reconstruction accuracy:

Differences in error between reconstruction methods were observed for RMSE (χ^2 = 55.073, p < 0.001), Jaccard index (χ^2 = 58.098, p < 0.001), and maximum error (χ^2 = 24.927, p < 0.001) (Figure 3 and 4). Median (IQR) errors were smaller in SSM reconstructions using nine landmarks (RMSE = 1.62 (0.35) mm, maximum error = 5.12 (1.63) mm) compared to isometric scaling (RMSE = 1.78 (0.62) mm, maximum error = 5.84 (2.62) mm, p < 0.001). Jaccard index was greater in SSM reconstructions from nine landmarks (0.824 (0.038)) compared to isometric scaling (0.792 (0.077), p < 0.001). The same pattern was observed between SSM reconstructions from fourteen landmarks (RMSE = 1.15 (0.29) mm, maximum error = (4.82 (1.26) mm, Jaccard index = 0.833 (0.034)) compared to isometric scaling (p < 0.001). Differences between SSM reconstructions using nine and fourteen landmarks were also significant, where RMSE was 6.8% smaller and Jaccard index was 1.1% larger in reconstructions using fourteen landmarks (p < 0.001). A similar pattern of results was observed when comparing errors between reconstruction methods for specific regions of the tibia and fibula (see Supplementary Table S1). Shape errors tended to be larger in the fibula than the tibia, and in the proximal region.

Generalizability to older adults:

Differences between reconstruction methods for older adults were also observed for maximum error (F = 14.047, p = 0.004, η^2 = 0.609, Figure 5 and 6), and Jaccard index (F = 14.379, p = 0.004, η^2 = 0.615). Mean (SD) maximum errors were smaller in SSM reconstructions using nine landmarks (6.90 (1.00) mm) compared to isometric scaling (9.21 (2.36) mm, $p = 0.005$). Jaccard index was greater in SSM reconstructions from nine landmarks (0.769 (0.032)) compared to isometric scaling (0.672 (0.077), $p = 0.004$). The same pattern was observed between SSM reconstructions from fourteen landmarks (maximum error = 7.04 (1.03) mm, Jaccard index = 0.763 (0.037)) compared to isometric scaling ($p \leq 0.005$). Regional analysis indicated that differences in maximum error were driven by improvements in the proximal region of the tibia (Supplementary Table S1). No differences in error measures were observed in the fibula. Pairwise comparisons between SSM reconstructions and isometric scaling for RMSE for the tibia and fibula combined were not significant. RMSE at the proximal and distal regions of the tibia were smaller in SSM reconstructions when compared to isometric scaling ($p \leq 0.001$, Supplementary Table S1). None of the error measures were different between reconstructions from nine and fourteen landmarks.

Muscle moment arms

In the young adult group, the reconstruction method changed the maximum difference in moment arms, relative to the CT-based bones, of the semimembranosus, and the long and short head of biceps femoris ($\chi^2 \geq 14.244$, $p \leq 0.001$, Supplementary Figure S1); no differences in moment arms for other muscles originating from or inserting on the tibia-fibular complex were observed. SSM reconstructions had smaller differences in moment arms than isometric scaling for the biceps femoris long head (median (IQR): nine landmarks = 2.36 (1.90) mm, fourteen landmarks = 2.45 (2.08) mm, isometric scaling = 3.65 (2.90) mm, $p < 0.001$) and short head (median (IQR): nine landmarks = 3.01 (2.06) mm, fourteen landmarks = 2.74 (2.34) mm, isometric scaling = 3.99 (3.14) mm, $p \leq 0.001$). Differences for semimembranosus were smaller in reconstructions from fourteen landmarks (2.61 (1.46) mm) when compared to nine landmarks (3.09 (2.24) mm) and isometric scaling (3.47 (2.65) mm, $p < 0.001$). Moment arm differences in the older adults were larger than for the younger adults (4.76 – 8.33 mm vs 2.36 – 3.99 mm, respectively). No differences in muscle moment arms between reconstruction methods were observed for the older adults.

Discussion

The purpose of this study was to evaluate the accuracy of tibia-fibula reconstructions from anatomical bony landmarks using a SSM developed for a healthy young adult population, and quantify the subsequent effects on muscle moment arms. The secondary objectives were (1) to quantify sex differences in bone geometry within the young adult sample, and (2) to evaluate the generalizability of the developed SSM to older adults. SSM reconstructions reduced geometry errors and changed hamstring moment arms, when compared to isometric scaling. On average, females had slightly narrower proximal epiphyseal regions and less diaphyseal curvature. SSM reconstructions, isometric scaling, and muscle moment arms in older adults were less accurate compared to the young adults.

Reconstruction accuracy was better in the SSM-generated models compared to isometric scaling. The magnitude of the differences between SSM-generated reconstructions from nine and fourteen landmarks and isometrically scaled geometries in this study were somewhat small; RMSE was reduced by 0.16-0.27 mm (9-15%) and maximum error was reduced by 0.72-1.02 mm (12-17%) in SSM reconstructions compared to isometric scaling, depending on the number of anatomical landmarks. Nolte et al. (2020) observed a larger reduction in RMSE (0.99 mm, 26%) using SSM-based reconstruction with only one principal component compared to isometric scaling, and even greater reductions in error were observed when more principal components were used. Zhang et al (2016) observed a reduction in RMSE of 0.41 mm (11%) using SSM-based reconstruction from only three bony landmarks compared to linear scaling. The low reconstruction errors observed for isometric scaling in this study likely explains the smaller reductions in RMSE compared to previous work. Isometric scaling was more accurate in this study (RMSE = 1.78 mm) compared to Nolte et al. (2020) (RMSE = 3.87 mm) and Zhang et al. (2016) (RMSE = 3.63 mm). In fact, the RMSE for isometric scaling in this study was also lower than SSM-based tibia-fibula reconstruction errors observed by Nolte et al (2020) (2.88 mm) and Zhang et al. (2016) (3.22 mm). This could be explained, in part, by the small range of young participants used to create the SSM and evaluate reconstruction accuracy in this study, i.e., 18-24 years, compared to 15-92 years in Zhang et al. (2016) and 23-70 years in Nolte et al. (2020) - which likely included less geometric variability. The use of landmarks identified directly on the bone surface, as compared to skin markers may have also contributed to more accurate results when compared to isometric scaling in previous studies.

When the SSM based on the younger group was used to reconstruct tibia-fibula geometries for the older adult group (71-98 years), errors for both isometric scaling and SSM-based reconstructions were larger than errors for the young group. Isometric scaling of the average young adult tibia and fibula geometries consistently underestimated cross-sectional size throughout the length of the bones in older adults, which could be explained by periosteal expansion (Ruff & Hayes 1988; Jee 2001). SSM reconstructions were able to account for some of the variation, reducing the overestimation of cross-sectional size. Although SSM-based reconstruction errors were larger in the older group than the younger group, the SSM still provided 14% and 25% reductions in Jaccard index and maximum error, respectively, within the older group, indicating greater robustness for application to new populations when compared to isometric scaling of a generic geometry.

Geometry errors, particularly at locations affecting joint alignment and muscle attachment points, can substantially influence musculoskeletal model outcomes (Scheys et al. 2008; Gerus et al. 2013; Xu et al. 2020). RMSE and maximum errors were larger in the proximal epiphysis and metaphysis regions, where many muscles crossing the knee insert, than in the diaphysis of the tibia and fibula (Supplementary Tables S1 and S2). Moment arms of some muscles attaching in the proximal regions were different when using the isometrically scaled average geometry compared to the CT-based geometry, adding support to previous findings at the knee and hip (Scheys et al. 2008; Bahl et al. 2019). In young adults, the 9-17% reductions in geometric errors from SSM-based reconstruction resulted in 25-35% reductions in maximum moment arm differences for semimembranosus and biceps femoris long and short heads. Previous work has demonstrated substantial sensitivity of muscle forces and joint contact forces to geometry (image-based vs. generic models) and perturbations of muscle insertion points and moment arms on the order of ±1 cm (Carbone et al. 2012; Gerus et al. 2013; Xu et al. 2020). The absolute changes in muscle point and maximum moment arm differences between SSM and isometric scaling methods in this study were an order of magnitude smaller (≤1.3 mm). These differences may not translate to significant changes in muscle forces or joint contact forces, but this must be confirmed in future work. In the older adults, moment arm and muscle origin/insertion point differences relative to the CT-based model were larger: up to 8 mm and 22 mm, respectively (Supplementary Tables S3 and S4). However, moment arm differences were not changed between reconstruction methods. Although SSM-reconstruction improved geometric accuracy, it was not better than isometric scaling of a generic model for musculoskeletal modelling application in an outside population. A SSM including older adults in the training set would provide better results.

In addition to age, sex is a factor known to influence bone geometry. In this study, the scores for three principal components differentiated between sexes. While not statistically significant, likely due to the extremely conservative Bonferroni adjustment for eight comparisons, trends were observed in principal components 4 and 7. Shape differences were subtle, as these principal components accounted for very small percentages (3.15% and 1.19%) of variance in the SSM and there was overlap in the principal component scores (Figure 2). These results are consistent with the observations of Brzobohatá et al. (2016).

A limitation of this study is that landmarks were identified directly on the bone surface. In the young adult group, using only the 'palpable' landmarks slightly reduced accuracy when compared to reconstructions from all 14 landmarks, but this was still 4-12% better than isometric scaling. This illustrates the potential for a subset of landmarks that might be used to predict tibia-fibula geometry without the use medical imaging, perhaps using skin-mounted markers collected during a static motion capture trial. Of course, estimating the soft-tissue offset between skin mounted markers and bony landmarks and landmark placement errors may introduce additional uncertainty. Mean soft tissue offsets of 4.8-7.7 mm and skin marker placement inter-examiner precision of 11-20 mm have been reported for anatomical landmarks on the shank (Della Croce et al. 1999; Nolte et al. 2020). Methods have been proposed to reduce errors and improve reliability for skin marker placement (Osis et al. 2016; Hutchinson et al. 2018). Larger errors may be observed for markers on the tibia shaft, which would be placed using a measuring tape to identify 25, 50, and 75% positions along the tibial crest between the lateral malleolus and fibular head markers. An approach allowing axial movement of the tibial crest markers (Nolte et al. 2020) may reduce the effects of this source of error. Encouragingly, Nolte et al. (2020) observed small standard deviations (0.90-2.99 mm) in soft-tissue offsets for seven markers on the shank, six of which were the same or similar to landmarks used in this study. The authors reported that no differences in RMSE were observed between reconstructions from bone landmarks and skin markers digitised using an optical motion capture system, with or without soft-tissue offset corrections, when one or two principal components were used. This provides some confidence that the SSM developed in this study could be used to reconstruct tibia-fibula geometries using skin-mounted markers, although additional work is needed to determine the number of principal components that could be used and to quantify the model-specific reconstruction accuracy.

The training set used in this study to create the SSM, which was composed of young active adults with no musculoskeletal abnormalities, may limit the applicability of the model to clinical or paediatric populations. Ethnicity is also a factor influencing bone geometry (Mahfouz et al. 2012). Unfortunately, ethnicity information was not collected, although most participants appeared to be of Caucasian descent.

In conclusion, within a young physically active population, and using an average model specific to that population, isometric scaling provided predictions of tibia and fibula bone geometry with low error. The developed SSM produced estimated tibia and fibula geometries from bony landmarks with even greater accuracy. However, this only affected the moment arms of three muscles. Geometry errors were larger in the older adult group. Although SSM-based reconstruction using a model trained on young adults was able to account for some geometric variation in an outside population, it was not sufficiently robust to alter musculoskeletal model parameters compared to a scaled generic model.

Acknowledgements

This work was funded in part by the Natural Sciences and Engineering Research Council of Canada (RGPIN 01029-2015, 02404-2021 and CGS D – 534891 - 2019). The authors would like to thank Andrew S. Michalski, PhD, for collecting and providing access to the CT data collected on the older adults.

Declaration of Interest

The authors have no potential conflicts of interest to report.

References

Arnold EM, Ward SR, Lieber RL, Delp SL. 2010. A model of the lower limb for analysis of human movement. Ann Biomed Eng. 38(2):269–279. doi: 10.1007/s10439-009-9852-5

Audenaert EA, Pattyn C, Steenackers G, De Roeck J, Vandermeulen D, Claes P. 2019. Statistical Shape Modeling of Skeletal Anatomy for Sex Discrimination: Their Training Size, Sexual Dimorphism, and Asymmetry. Front Bioeng Biotechnol. 7:1–11. doi: 10.3389/fbioe.2019.00302

Bahl JS, Zhang J, Killen BA, Taylor M, Solomon LB, Arnold JB, Lloyd DG, Besier TF, Thewlis D. 2019. Statistical shape modelling versus linear scaling: Effects on predictions of hip joint centre location and muscle moment arms in people with hip osteoarthritis. J Biomech. 85:164-172. doi: 10.1016/j.jbiomech.2019.01.031

Brzobohatá H, Krajíček V, Horák Z, Velemínská J. 2015. Sex classification using the three-dimensional tibia form or shape including population specificity approach. J Forensic Sci. 60(1):29–40. doi: 10.1111/1556-4029.12641

Brzobohatá H, Krajíèek V, Horák Z, Velemínská J. 2016. Sexual dimorphism of the human tibia through time: Insights into shape variation using a surface-based approach. PLoS One. 11(11):e0166461. doi: 10.1371/journal.pone.0166461

Carbone V, van der Krogt MM, Koopman HFJM, Verdonschot N. 2012. Sensitivity of subject-specific models to errors in musculo-skeletal geometry. J Biomech. 45(14):2476–2480. doi: 10.1016/j.jbiomech.2012.06.026

Clouthier AL, Smith CR, Vignos MF, Thelen DG, Deluzio KJ, Rainbow MJ. 2019. The effect of articular geometry features identified using statistical shape modelling on knee biomechanics. Med Eng Phys. 66:47-55. doi: 10.1016/j.medengphy.2019.02.009

Della Croce U, Cappozzo A, Kerrigan DC. 1999. Pelvis and lower limb anatomical landmark calibration precision and its propagation to bone geometry and joint angles. Med Biol Eng Comput. 37(2):155–161. doi: 10.1007/BF02513282

Delp SL, Loan JP, Hoy MG, Zajac FE, Topp EL, Rosen JM. 1990. An Interactive Graphics-Based Model of the Lower Extremity to Study Orthopaedic Surgical Procedures. IEEE Trans Biomed Eng. 37(8):757-767. doi: 10.1109/10.102791

Ding Z, Tsang CK, Nolte D, Kedgley AE, Bull AM. 2019. Improving musculoskeletal model scaling using an anatomical atlas: the importance of gender and anthropometric similarity to quantify joint reaction forces. IEEE Trans Biomed Eng. 66(12):3444-3456. doi: 10.1109/TBME.2019.2905956

Edwards WB, Schnitzer TJ, Troy KL. 2013. Torsional stiffness and strength of the proximal tibia are better predicted by finite element models than DXA or QCT. J Biomech. 46(10):1655–1662. doi: 10.1016/j.jbiomech.2013.04.016

Feldman S, Capozza RF, Mortarino PA, Reina PS, Ferretti JL, Rittweger J, Cointry GR. 2012. Site and sex effects on tibia structure in distance runners and untrained people. Med Sci Sports Exerc. 44(8):1580–1588. doi: 10.1249/MSS.0b013e31824e10b6

Gerus P, Sartori M, Besier TF, Fregly BJ, Delp SL, Banks SA, Pandy MG, D'Lima DD, Lloyd DG. 2013. Subject-specific knee joint geometry improves predictions of medial tibiofemoral contact forces. J Biomech. 46(16):2778–2786. doi: 10.1016/j.jbiomech.2013.09.005

Grant TM, Diamond LE, Pizzolato C, Killen BA, Devaprakash D, Kelly L, Maharaj JN, Saxby DJ. 2020. Development and validation of statistical shape models of the primary functional bone segments of the foot. PeerJ. 8:e8397. doi: 10.7717/peerj.8397

Hutchinson L, Schwartz JB, Morton AM, Davis IS, Deluzio KJ, Rainbow MJ. 2018. Operator Bias Errors Are Reduced Using Standing Marker Alignment Device for Repeated Visit Studies. J Biomech Eng. 140(4):1–7. doi: 10.1115/1.4038358

Jee WSS. 2001. Integrated bone tissue physiology: Anatomy and physiology. In: Cowin SC, editor. Bone Mechanics Handbook. 2nd Ed. Boca Raton (FL): CRC Press; p. 1-68

Li P, Tsai TY, Li JS, Zhang Y, Kwon YM, Rubash HE, Li G. 2014. Morphological measurement of the knee: Race and sex effects. Acta Orthop Belg. 80(2):260-268.

Mahfouz M, Fatah EEHA, Bowers LS, Scuderi G. 2012. Three-dimensional morphology of the knee reveals ethnic differences. Clin Orthop Relat Res. 470(1):172–185. doi: 10.1007/s11999-011-2089-2

Mei L, Figl M, Rueckert D, Darzi A, Edwards P. 2008. Sample sufficiency and number of modes to retain in statistical shape modelling. In: Metaxas D, Axel L, Fichtinger G, Szekely G, editors. Lect Notes Comput Sci. Berlin: Springer; p. 425–433.

Myronenko A, Song X. 2010. Point set registration: Coherent point drift. IEEE Trans Pattern Anal Mach Intell. 32(12):2262–2275. doi: 10.1109/TPAMI.2010.46

Nolte D, Ko ST, Bull AMJ, Kedgley AE. 2020. Reconstruction of the lower limb bones from digitised anatomical landmarks using statistical shape modelling. Gait Posture. 77:269–275. doi: 10.1016/j.gaitpost.2020.02.010

Osis ST, Hettinga BA, Macdonald S, Ferber R. 2016. Effects of simulated marker placement deviations on running kinematics and evaluation of a morphometric-based placement feedback method. PLoS One. 11(1):1–13. doi: 10.1371/journal.pone.0147111

Real R, Vargas JM. 1996. The Probabilistic Basis of Jaccard's Index of Similarity. Syst Biol. 45(3):380–385. doi: 10.1093/sysbio/45.3.380

Ruff CB, Hayes WC. 1988. Sex differences in age-related remodeling of the femur and tibia. J Orthop Res. 6(6):886–896. doi: 10.1002/jor.1100060613

Scheys L, Spaepen A, Suetens P, Jonkers I. 2008. Calculated moment-arm and muscle-tendon lengths during gait differ substantially using MR based versus rescaled generic lower-limb musculoskeletal models. Gait Posture. 28(4):640–648. doi: 10.1016/j.gaitpost.2008.04.010

Sherk VD, Bemben DA, Bemben MG, Anderson MA. 2012. Age and sex differences in tibia morphology in healthy adult Caucasians. Bone. 50(6):1324–1331. doi: 10.1016/j.bone.2012.03.005

Smock AJ, Hughes JM, Popp KL, Wetzsteon RJ, Stovitz SD, Kaufman BC, Kurzer MS, Petit MA. 2009. Bone volumetric density, geometry, and strength in female and male collegiate runners. Med Sci Sports Exerc. 41(11):2026-2032. doi: 10.1249/MSS.0b013e3181a7a5a2

Soucie JM, Wang C, Forsyth A, Funk S, Denny M, Roach KE, Boone D. 2011. Range of motion measurements: Reference values and a database for comparison studies. Haemophilia. 17(3):500–507. doi: 10.1111/j.1365-2516.2010.02399.x

Stevens SD, Vidarsdóttir US. 2008. Morphological changes in the shape of the non-pathological bony knee joint with age: A morphometric analysis of the distal femur and proximal Tibia in three populations of known age at death. Int J Osteoarchaeol. 18(4):352–371. doi: 10.1002/oa.954

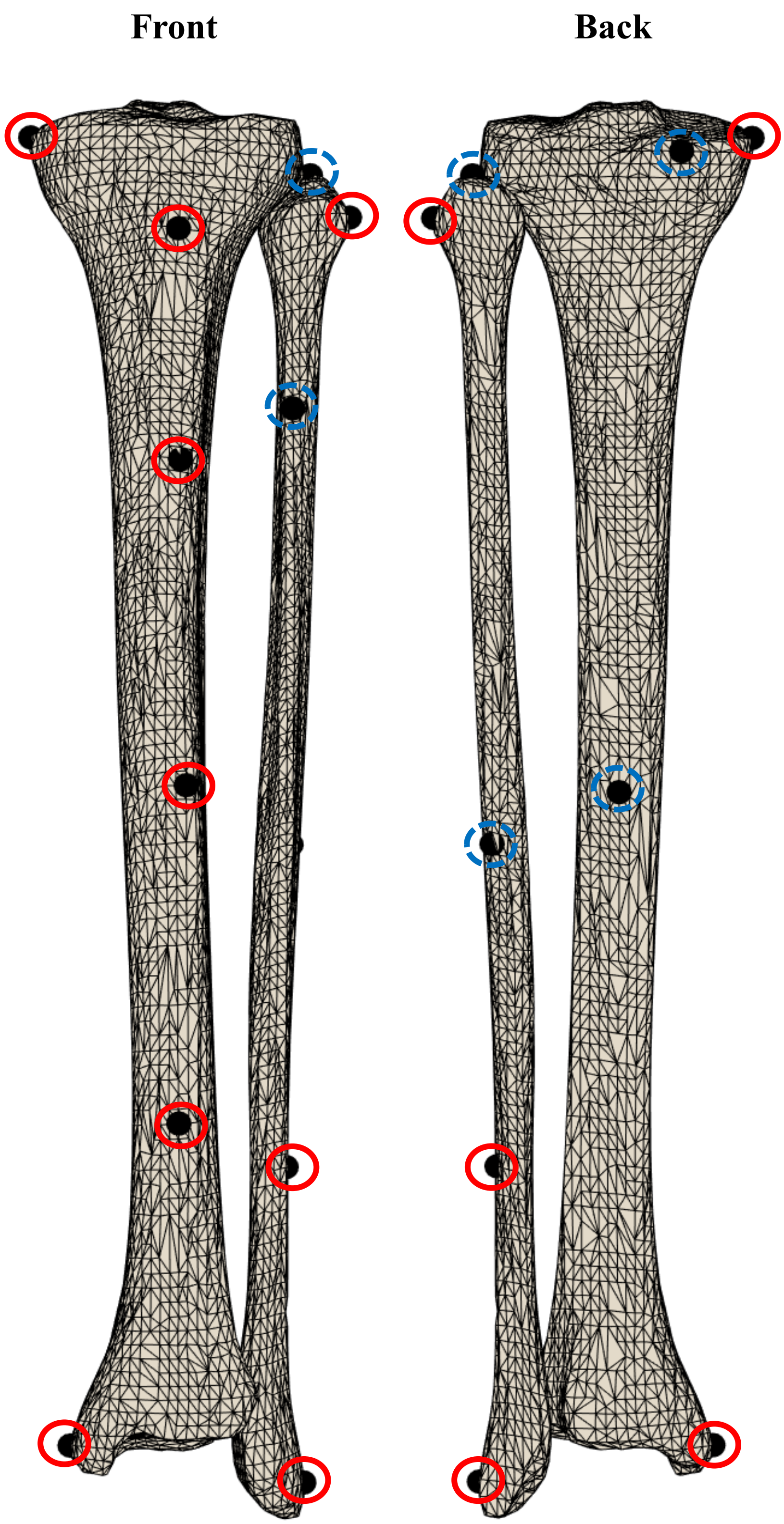
Tümer N, Arbabi V, Gielis WP, de Jong PA, Weinans H, Tuijthof GJM, Zadpoor AA. 2019. Three-dimensional analysis of shape variations and symmetry of the fibula, tibia, calcaneus and talus. J Anat. 234(1):132-144. doi: 10.1111/joa.12900

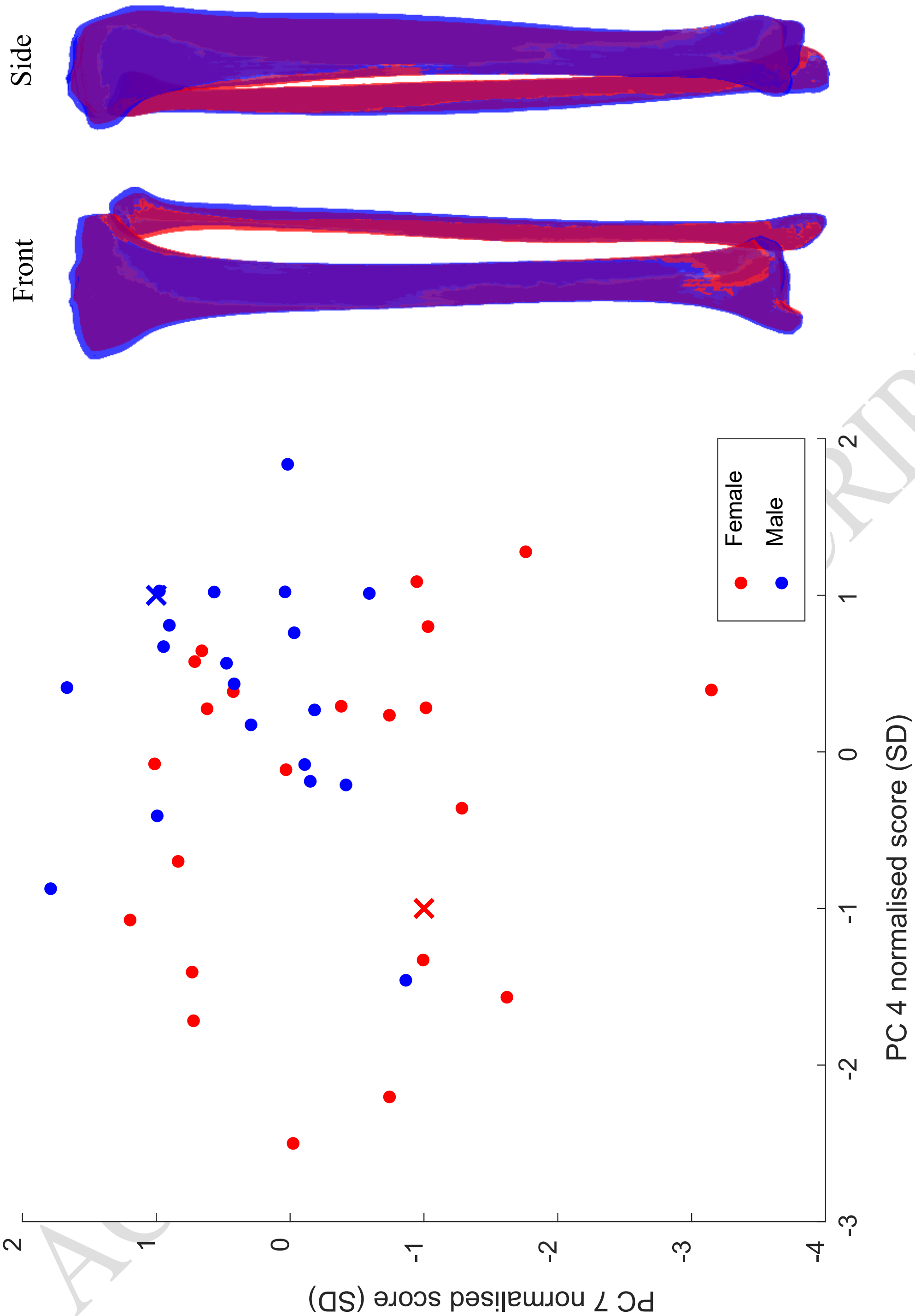
Xu C, Reifman J, Baggaley M, Edwards WB, Unnikrishnan G. 2020. Individual Differences in Women during Walking Affect Tibial Response to Load Carriage: The Importance of Individualized Musculoskeletal Finite-Element Models. IEEE Trans Biomed Eng. 67(2):545-555. doi: 10.1109/TBME.2019.2917415

Zhang J, Fernandez J, Hislop-Jambrich J, Besier TF. 2016. Lower limb estimation from sparse landmarks using an articulated shape model. J Biomech. 49(16):3875-3881. doi: 10.1016/j.jbiomech.2016.10.021

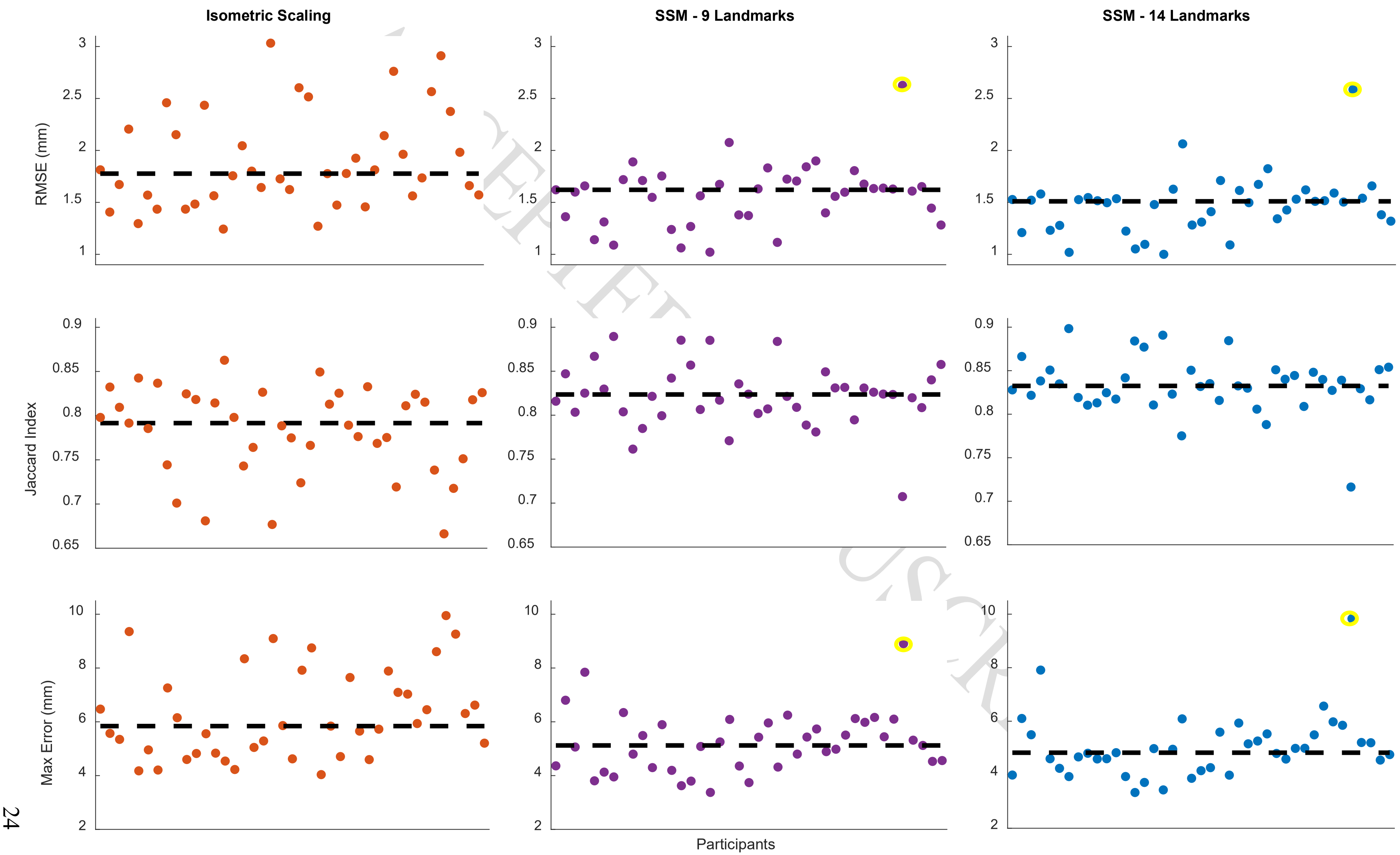
Table 1. Percent of total variance explained by principal components 1-8.

Principal component	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Variance explained (% of total)	79.8	4.7	3.6	3.1	1.7	1.4	1.2	0.7
Cumulative variance explained (% of total)	79.8	84.5	88.1	91.2	92.9	94.3	95.5	96.2

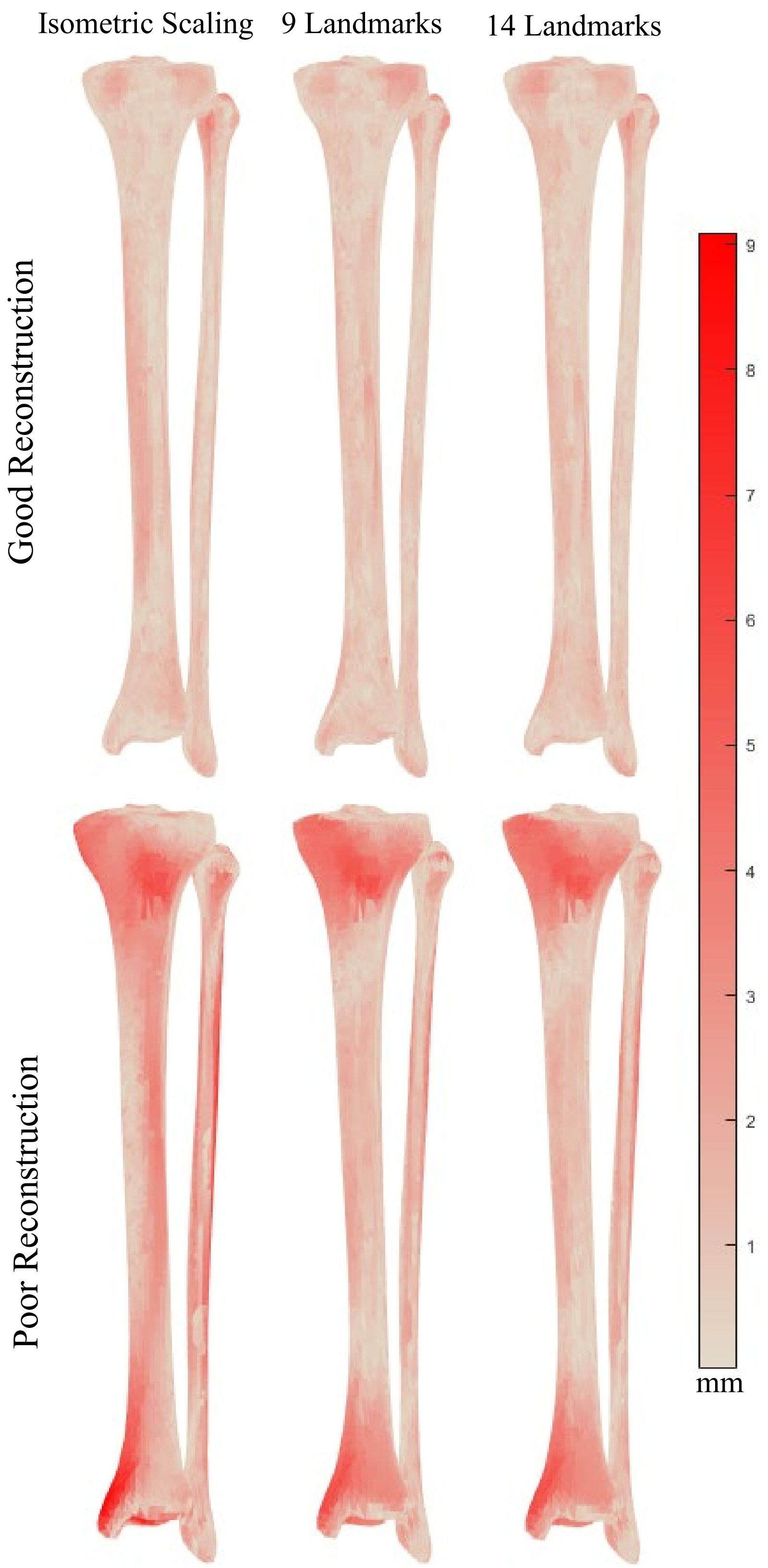

Front
Back



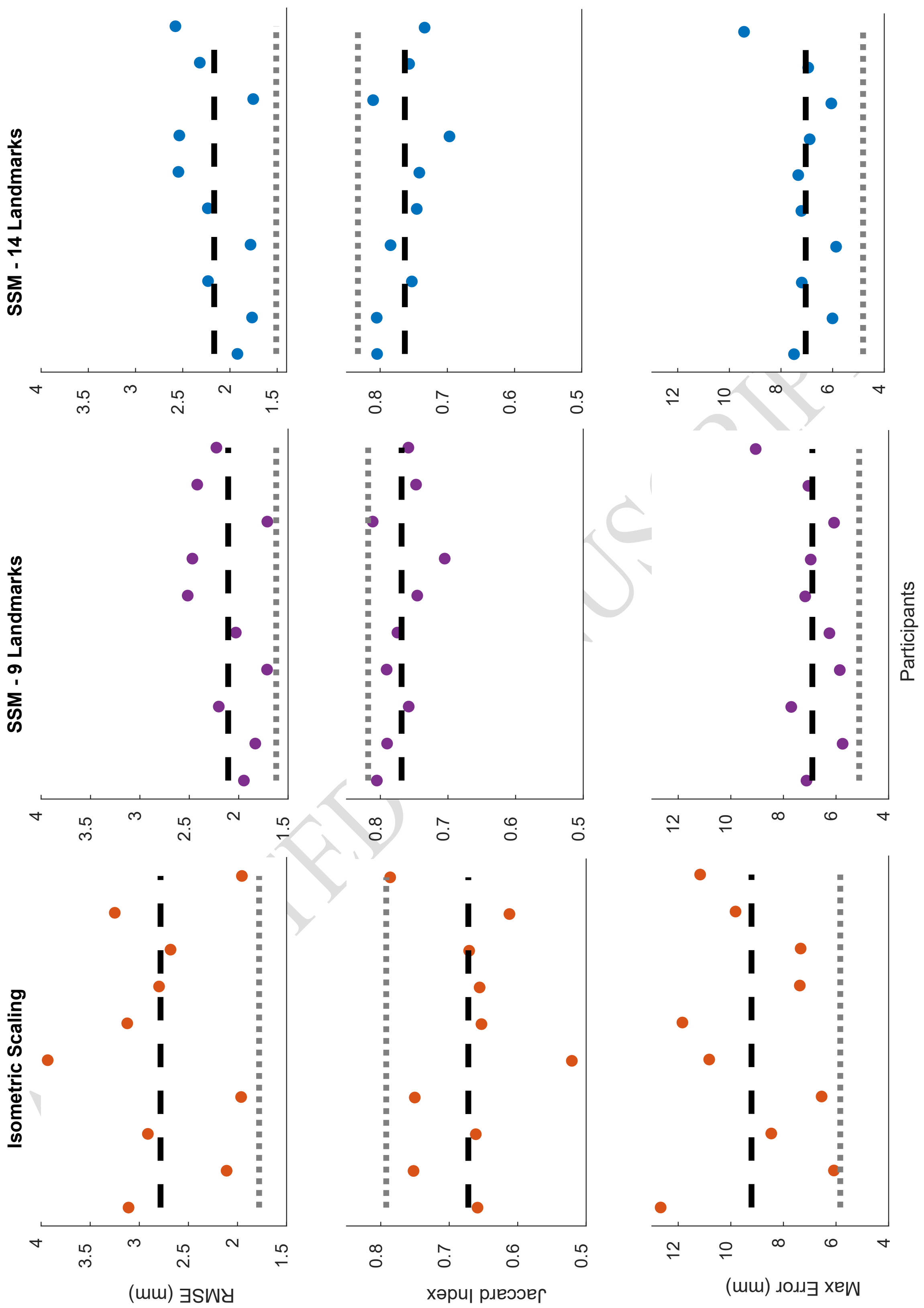
Side
Front
Female
Male
PC 4 normalised score (SD)
PC 7 normalised score (SD)



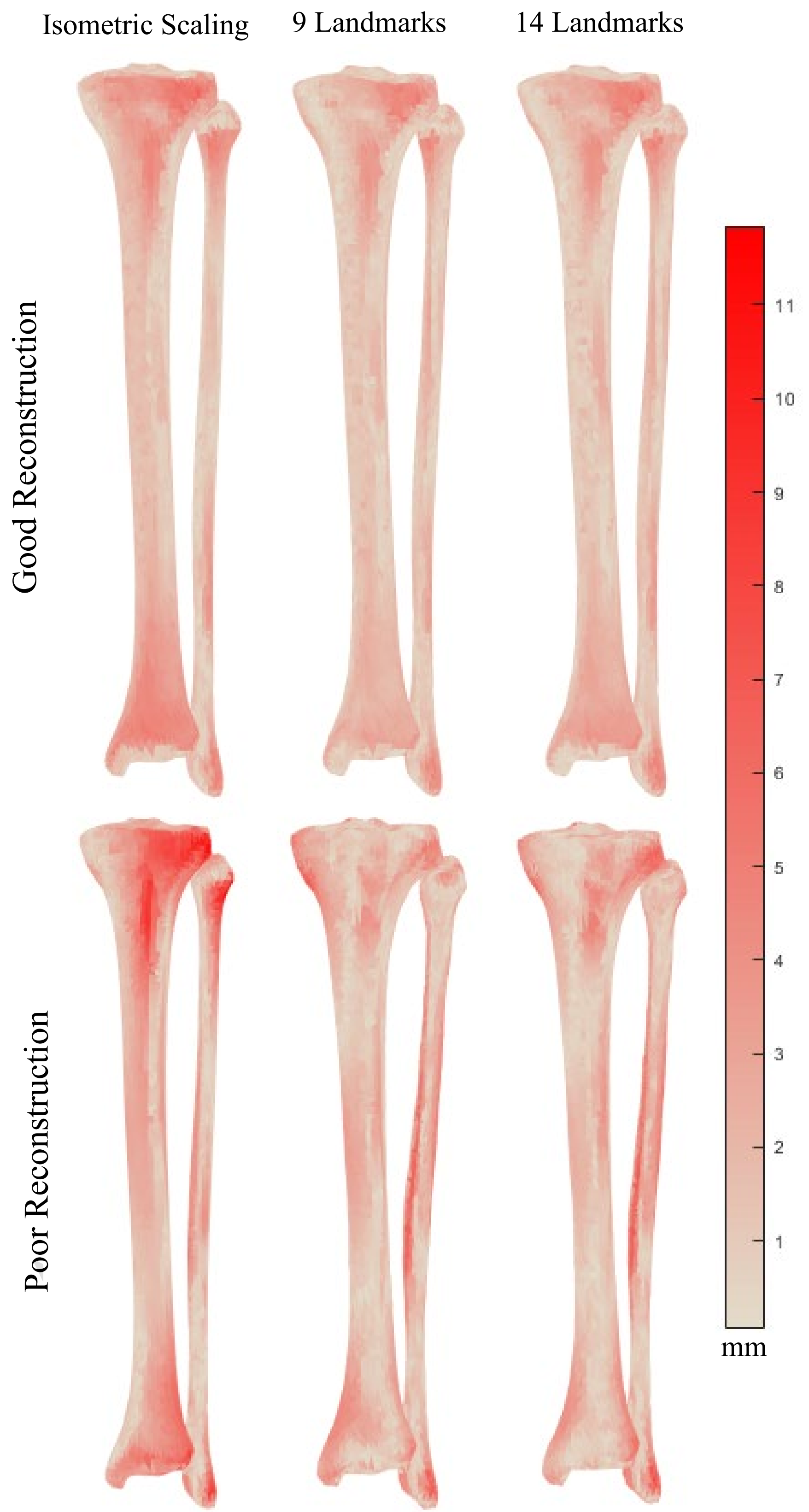
Isometric Scaling
SSM - 9 Landmarks
SSM - 14 Landmarks
RMSE (mm)
Jaccard Index
Max Error (mm)
Participants
3
2.5
2
1.5
1
0.9
0.85
0.8
0.75
0.7
0.65
10
8
6
4
2



Isometric Scaling
9 Landmarks
14 Landmarks
Good Reconstruction
Poor Reconstruction
9
8
7
6
5
4
3
2
1
mm



Isometric Scaling
SSM - 9 Landmarks
SSM - 14 Landmarks
RMSE (mm)
Jaccard Index
Max Error (mm)
Participants



Isometric Scaling
9 Landmarks
14 Landmarks
Good Reconstruction
Poor Reconstruction
11
10
9
8
7
6
5
4
3
2
1
mm

520

521

522

Figure 1. Landmarks used for reconstructions. The subset of nine palpable landmarks are circled in red. The five non-palpable landmarks are circled in blue (dashed line).

Figure 2. (Left) Scatterplot of principal components 4 and 7 normalised scores. Red circles (female) and blue squares (male) represent individual participants. The x's represent the average shape + (blue) or – (red) 1 standard deviation (SD) for both principal components (PC) 4 and 7. (Left) Representations of the tibia-fibula complex geometry of the blue and red x's. Sex differences in shape were very subtle. ± 1 SD was larger than mean normalised scores for females (PC4 = -0.31, PC7 = -0.31) and males (PC4 = 0.36, PC7 = 0.35) and was used to more easily visualise differences.

Figure 3. Errors and volume similarity of tibia and fibula geometries predicted using isometric scaling or SSM-landmark reconstruction methods compared to CT data. Dots represent individual participants. The highlighted dots are an outlier. The dashed line represents the median. Differences between SSM-landmark methods and isometric scaling were significant for all three measures. Differences between 9 and 14 landmark reconstructions were significant for RMSE and Jaccard index.

Figure 4. Good (top, participant 14) and poor (bottom, participant 19) reconstructions for isometric scaling and SSM-landmark methods.

Figure 5. Errors and volume similarity of tibia and fibula geometries for older individuals predicted using isometric scaling or SSM-landmark reconstruction methods compared to CT data. Dots represent individual participants. The black dashed line represents the mean. Differences between SSM-landmark methods and isometric scaling were significant for Jaccard index and maximum error. Errors were not different between 9 and 14 landmark reconstructions. The dotted grey line represents the median of the young adult group.

Figure 6. Good (top, participant 4) and poor (bottom, participant 6) reconstructions of older participants for isometric scaling and SSM-landmark methods.