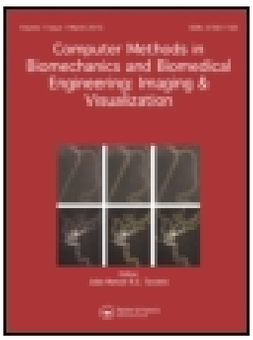

Taylor & Francis
Taylor & Francis Group

Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization

ISSN: 2168-1163 (Print) 2168-1171 (Online) Journal homepage: http://www.tandfonline.com/loi/tciv20

Morphological errors in 3D bone models of the distal femur and proximal tibia generated from magnetic resonance imaging and computed tomography determined using two registration methods

Valentina Campanelli, Stephen M. Howell & Maury L. Hull

To cite this article: Valentina Campanelli, Stephen M. Howell & Maury L. Hull (2019): Morphological errors in 3D bone models of the distal femur and proximal tibia generated from magnetic resonance imaging and computed tomography determined using two registration methods, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, DOI: 10.1080/21681163.2018.1559101

To link to this article: https://doi.org/10.1080/21681163.2018.1559101

Published online: 02 Jan 2019.

Submit your article to this journal

View Crossmark data

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=tciv20

COMPUTER METHODS IN BIOMECHANICS AND BIOMEDICAL ENGINEERING: IMAGING & VISUALIZATION
https://doi.org/10.1080/21681163.2018.1559101

Taylor & Francis
Taylor & Francis Group

Check for updates

Morphological errors in 3D bone models of the distal femur and proximal tibia generated from magnetic resonance imaging and computed tomography determined using two registration methods

Valentina Campanelli[a], Stephen M. Howell[b] and Maury L. Hull[a,b,c]

[a]Department of Mechanical Engineering, University of California, Davis, CA, USA; [b]Department of Biomedical Engineering, University of California, Davis, CA, USA; [c]Department of Orthopaedic Surgery, University of California, Davis, CA, USA

ABSTRACT

Three dimensional models of the femur and tibia from MRI and CT are used in various studies of knee biomechanics but morphological errors occur. The objectives were to determine the morphological errors in 3D bone models from MRI and CT using clinically-adopted imaging protocols, to determine whether these errors are affected by the method of registration to a gold-standard 3D model, and to determine whether the slice thickness of CT affects errors. CT and MRI scans of 13 human cadaveric knees were performed and 3D bone models of the distal femur and proximal tibia were compared to a gold-standard 3D bone model generated by laser scanning. Registration of the models from MRI and CT to those from laser scanning was performed using the iterative closest point (ICP) algorithm and fiducial markers. The key findings were that 1) 3D bone models generated with CT yield submillimeter accuracy while 3D models generated with MRI yield accuracy worse than 1 mm; 2) CT bone models slightly overestimate the bone morphology, while the MRI bone models substantially underestimate it; 3) morphological errors of 3D bone models from 0.625 mm and 1.25 mm slice thickness CT scans are not different; and 4) ICP registration underestimates the morphological errors compared to registration with fiducial markers.

ARTICLE HISTORY
Received 29 March 2018
Accepted 11 December 2018

KEYWORDS
3D bone models; total knee arthroplasty; computed tomography; magnetic resonance imaging; medical imaging; slice thickness; knee; bone morphology; accuracy; laser scanner

Introduction

Subject-specific three-dimensional (3D) bone models of the distal femur and proximal tibia are used in various applications related to knee biomechanics. Example applications include computational models (Donahue et al. 2002, Kang et al. 2017), study of knee kinematics and tibial contact kinematics using fluoroscopy (Bellemans et al. 2002, Komistek et al. 2003, Fregly et al. 2005), design of implants used in total knee replacement (Kwak et al. 2007, Cheng et al. 2009), presurgical planning of total knee replacement (My Knee, Truematch Personalized Solutions) and design and manufacturing of patient-specific cutting guides for total knee replacement surgery (My Knee, Truematch Personalized Solutions). Some of these applications require 3D bone models with submillimeter accuracy. One such application is the study of kinematics and tibial contact kinematics using fluoroscopy (Komistek et al. 2003, Fregly et al. 2005). If the 3D bone model is not accurate, then the 3D bone model-to-2D image registration process, which relies on matching projections of the 3D bone models to the outlines of the bones in the images, is prone to error (Kaptein et al. 2003, Moro-Oka et al. 2007). Another such application is design and manufacturing of patient-specific cutting guides for total knee replacement surgery. Errors in making the distal and posterior femoral resections as small as 0.5 mm cause significant increases in tibial force imbalance (Riley et al. 2018a, 2018b) which have been related to poorer patient-reported outcomes (Gustke et al. 2014).

The 3D bone models are often generated with computed tomography (CT) or magnetic resonance imaging (MRI), using low-resolution imaging protocols to reduce ionising radiation (CT) and scanning time (MRI). However, various sources of error are inherent to each imaging modality and protocol (e.g. noise) and to the post processing of the images (e.g. segmentation). Accordingly, morphological errors necessarily occur in 3D bone models generated from CT and MRI as demonstrated by studies comparing these models to a gold-standard 3D bone model (Gelaude et al. 2008, Rathnayaka et al. 2012, Van Den Broeck et al. 2014). However, because a number of factors can introduce artefacts into the determination of morphological errors, it is important to determine these errors using methods that minimise or eliminate artefacts.

One factor which can introduce artefacts in the determination of morphological errors is the gold-standard bone model per se if it is not accurate. Because gold-standard 3D bone models are often generated using a surface scanning modality (e.g. laser scanner or a contacting scanner), one requirement for accuracy is a method to remove the cartilage and soft tissue remains from the articular surfaces of the distal femur and proximal tibia without affecting the size of the bone, a process known as maceration. The maceration techniques used in previous studies include the immersion of the

CONTACT Valentina Campanelli valentina.campanelli@gmail.com Department of Mechanical Engineering, University of California, One Shields Avenue, Davis, CA 95616, USA

© 2018 Informa UK Limited, trading as Taylor & Francis Group

specimen in sodium hypochlorate (Donahue et al. 2002), or acetone (Van Den Broeck et al. 2014), and/or boiling (Gelaude et al. 2008, Van Den Broeck et al. 2014) which have been found to reduce the overall dimension of the bone by up to 0.2–0.5 mm (Gelaude et al. 2008, Van Den Broeck et al. 2014) thus affecting the accuracy of the gold-standard bone model. Bone cleaning via dermestid beetles has been considered the gold-standard among maceration techniques because it best preserves the size, shape, colour and mechanical properties of the bone (Steadman et al. 2006).

A second requirement in generating an accurate gold-standard 3D model is that the accuracy of the surface scanning modality should be preferably an order of magnitude better than the error to be detected. If morphological errors of MRI and CT 3D models range between 0.5 mm to 1 mm, then the scanner must be accurate within 50–100 μm. Previous studies only report laser scanner errors declared by the manufacturer (Gelaude et al. 2008, Rathnayaka et al. 2012, Van Den Broeck et al. 2014), which are based on scanning objects with simple shapes (e.g. gage blocks) and can be considerably lower than the error when scanning a freeform surface like a bone (Campanelli et al. 2016).

In addition to generating accurate gold-standard bone models, another factor which can introduce artefact is the method used to register the gold-standard bone model to the CT and MRI bone models. A popular method of 3D-3D registration is the iterative closest point algorithm (ICP) (Besl and Mckay 1992) because it is implemented in commercially available software. However, using ICP to register two models may result in the underestimation of morphological errors because ICP minimises the difference in the shapes of the two models. A more involved but also more accurate method of 3D-3D registration is based on fiducial markers implanted in the bone, because the resulting transformation matrix is independent from the intrinsic morphology of the models under study.

Previous studies are limited for a variety of reasons. One reason is that some were subject to one or more of the artefacts noted above (Gelaude et al. 2008, Rathnayaka et al. 2011, 2012, Van Den Broeck et al. 2014). Other reasons are that some used high-resolution imaging protocols that were not clinically adopted (Rathnayaka et al. 2012), did not use a gold-standard 3D model for comparison (Moro-Oka et al. 2007, White et al. 2008), and analysed the morphological errors for only the tibia and not the femur (Van Den Broeck et al. 2014).

The objectives of the present study were threefold. One was to determine the morphological errors of 3D bone models of human cadaveric specimens generated with MRI and CT using clinically adopted imaging protocols and a methodology that overcomes the limitations noted above. The second objective was to determine whether morphological errors are affected by the method of registration using both ICP and fiducial markers. A final objective was to determine whether the slice thickness of CT affected morphological errors. The latter objective was of interest because some clinically adopted imaging protocols use thicker slices to reduce the exposure to radiation.

Methods

A total of 13 unpaired fresh-frozen human cadaveric knees (average age: 81.2 years, age range: 63–101 years, 10 females and 3 males), each free of radiographic signs of degenerative joint disease, were included in the study. The knees were amputated mid-femur and mid-tibia, and the proximal part of the femur specimen and the distal part of the tibia specimen were potted using methylmethacrylate (MMA, Figure 1).

Fiducial markers were fabricated using a 3D printer (Objet Connex 260V, Stratasys, USA). Each marker consisted of a hollow sphere within a sphere 28 mm in outside diameter (Figure 2). Six to seven fiducial markers were attached to each of the femurs and tibias. Specifically, 4–5 markers were attached to the potting cup using a nylon-threaded stud and two-part resin epoxy glue, while the remaining two markers were attached directly to the bone using nylon-threaded studs and MMA (Figure 1).

Each knee was scanned with MRI and CT using clinically adopted imaging protocols for preoperative planning for patient-specific instruments (MRI) and robotic-surgery (CT) for total knee arthroplasty. For each specimen, two CT scans were performed with a 32-slice CT scanner (GE LightSpeed) using two imaging protocols that differed only in the slice thickness used, 0.625 mm and 1.25 mm. These two CT protocols were used as they are both commonly used for preoperative planning in TKA. The CT imaging protocol included 120 kVp, smart mA, no slice gap/overlap, a 512 × 512 image matrix, and a pixel size of 0.39 mm. One MRI scan was performed on each knee with a 1.5 T MRI scanner (GE, Genesis Signa) and a knee coil using a fast time of flight spoiled gradient recalled (FAST TOF SPGR) sequence and sagittal slices

Figure 1. Knee specimen after potting and application of the fiducial markers.

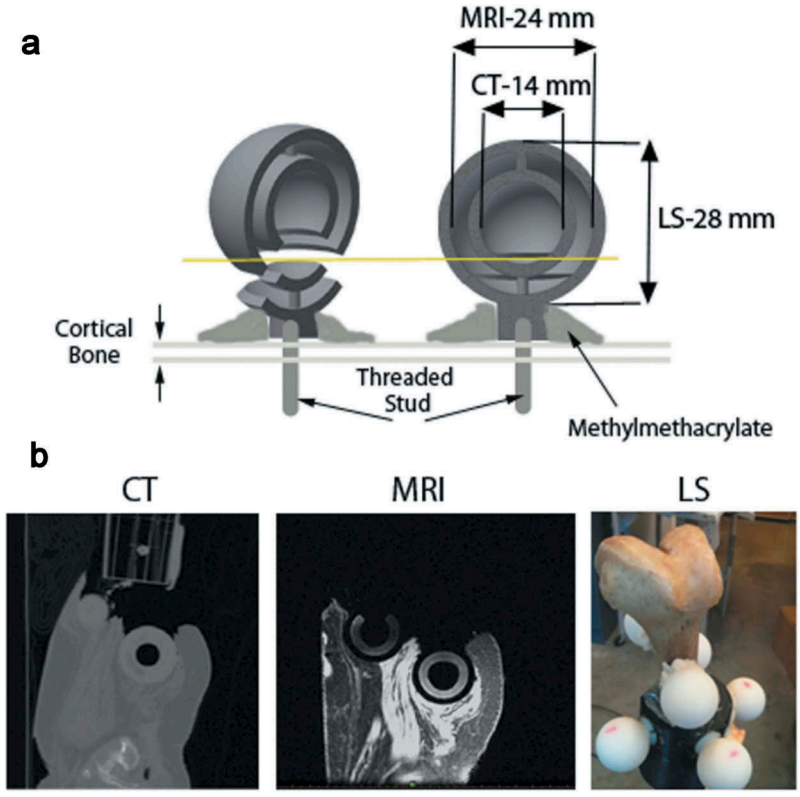

Figure 2. Schematic of the fiducial marker design and attachment to the bone. The design of the fiducial marker includes three structures (Figure 2(a)). The innermost structure is a hollow sphere 14 mm in diameter, which was designed to be clearly seen in CT images (Figure 2(b), left). The second structure is an annulus filled with 3D printing MRI-opaque support material with an outer diameter of 24 mm which was designed to be clearly seen in MRI images (Figure 2(b), middle). The outer surface of the marker was designed to be scanned by a 3D surface scanner (Figure 2(b), right). To remove the support material from the inner sphere, the marker assembly was designed in two parts, the sphere and the stem, which were glued together after prototyping and cleaning. The stem was tapped so that a threaded nylon stud could be inserted into the fiducial marker and be connected to the bone through methylmethacrylate or to the cup through epoxy. Figure 2b shows how the different structures of the markers are visualised in each imaging modality. The diameter of each structure was chosen so that the inner sphere could be visualised and segmented in at least 11 CT slices at 1.25 mm thickness (22 slices for a CT 0.625 mm), and so that the annulus could be visualised and segmented into at least 12 MRI slices at 2 mm slice thickness.

with TR = 17 s, TE = 4 s, flip angle = 25°, 2 mm slice thickness, 1 mm spacing between slices, a 256 × 256 image matrix interpolated to 512 × 512 pixels, pixel size = 0.78 mm interpolated to 0.39 mm. The MRI scanning plane was set approximately perpendicular to the flexion-extension axis of the knee. Because the fiducial markers located farther from the joint were outside the volume of the knee coil, two additional MRI scans (one for the femur and one for the tibia) were performed immediately following the MRI scan of the knee to scan all fiducial markers in each bone using the same protocol described above, but using a body coil rather than a knee coil. Care was taken to position the knee in the knee coil so that no movement of the specimen occurred between the MRI of the knee and the two successive MRI scans of the femur and tibia by using foam objects and towels to fill in the gap between the knee coil and the knee. Only 11 out of 13 tibias were included in the analysis as two fiducial markers broke in each specimen right before laser scanning. Also, one femur was excluded from the MRI analysis as two fiducial markers appeared to be highly distorted in the MRI image.

After the MRI and CT scans were performed, the knee specimen was grossly dissected to remove the soft tissues, disarticulated, and macerated with dermestid beetles to remove remaining cartilage and soft tissues. The maceration via dermestid beetles lasted on average 1 month for each bone (four colonies of beetles were used in parallel), during which the bones were water-sprayed every 2–3 days to avoid desiccation. Finally, the cleaned bones were laser scanned using a high-accuracy 3D laser scanner to generate a gold-standard model (GS) of the distal femur and proximal tibia surface with about 1–1.5 million triangles (Metrascan 3D, Creaform, Canada). The Metrascan 3D laser scanner captured the entire 3D anatomy of the bone with only one scan without the need of performing multiple scans and registering the scans in post-processing to obtain the entire 3D anatomy. The repeatability error (root mean square deviation) in scanning a human distal femur five times was 87 microns as determined by a pairwise comparison of the five 3D models (Campanelli et al. 2016), while the overall error (RMSE) reported by the manufacturer in scanning reference objects is 85 microns. This error was considered acceptable to detect 0.5–1 mm errors in MRI and CT bone models, as required for this study.

The MRI and CT images of each knee specimen were segmented using the automatic tools in Mimics® v20.0 (Materialise, Belgium) and refined manually. The CT images were segmented using a combination of thresholding and manual segmentation. The MRI images were first segmented using Smart Expand,

which is a region-growing algorithm implemented by Mimics that segments the region of interest (ROI) given a few initial segmented slices. After this initial step, manual segmentation was still used extensively to refine the selected ROI. Overall, the segmentation of the femur and tibia took 5 h for the MR images (about 250 images) and 7 h for the thinner-slice CT images (about 500 images for the 0.625 mm slice thickness data set). The fiducial markers were also segmented with Mimics using a combination of thresholding and manual segmentation for a total of 546 segmented markers (14 markers per specimen x 3 scans per specimen x 13 specimens).

Mimics was used to reconstruct 3D models of the femur, tibia and fiducial markers using a variation of the classic 'marching cubes algorithm' (Lorensen and Cline 1987). The following reconstruction settings were used: interpolation method 'gray value', preferred 'accuracy', shell reduction to 1, no matrix reduction applied, and a smoothing factor of 0.5 using 7 iterations. Hence, four femoral and four tibial 3D bone models (MRI—2 mm, CT—0.625 mm, CT—1.25 mm, GS) were generated for each knee (Figure 3) and each was exported in stereolithographic format and furtherly processed to remove any remaining inner surfaces inside the models using a commercially available software (Geomagic®, 3D Systems, Cary, NC).

The MRI and CT bone models of the femur and tibia were each registered to the gold-standard models using two methods (Figure 4). In the first method, the 3D models were registered using the centres of the spheres generated through a best-fit of the 3D models of the 6–7 fiducial markers using a least-squares fitting method implemented in MATLAB® R2017b (Mathworks, Natick, Massachusetts) (Arun et al. 1987). Before this step, a custom MATLAB routine was written to identify the best subset of fiducial markers to use for the registration for each specimen and each imaging modality. The main goal of this routine was to exclude from the registration those marker centres that were largely misidentified because of either image distortion for MRI, 3D printing error, or improper removal of the 3D printing material inside the fiducial markers for CT. The MATLAB routine registered the centres of the fiducial markers of the GS model to the centres of the fiducial markers of the CT model using all possible subsets of markers up to a minimum of four marker centres. For each subset, the fiducial registration error was computed as the root mean square distance between each pair of GS-CT paired markers and the winning subset used for registration was the one having the lowest fiducial registration error. The same procedure was used to identify the best marker subset for registration of the MRI 3D model.

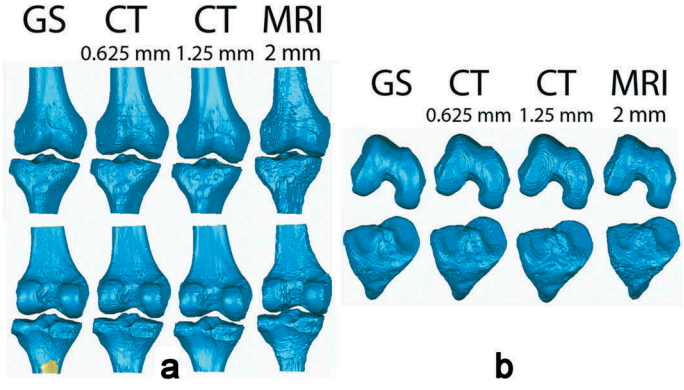

Figure 3. Coronal (a) and axial views (b) of 3D models of the distal femur and proximal tibia generated from MRI (one protocol), CT (two protocols, 0.625 and 1.25 mm slice thickness), and laser scanning (GS) for a representative knee specimen. The higher resolution of the GS femur and tibia models can clearly be seen in these views. The quality of the MRI model is inferior to the quality of the CT and GS models and the CT 1.25 mm model has a lower resolution than the CT 0.625 mm model, especially at the edges of the femur trochlea.

In the second method, the 3D models were registered using the ICP algorithm implemented in Geomagic 2015.3.1. The morphological errors of the MRI and CT 3D bone models were determined through a direct comparison with the GS 3D bone model by calculating the Euclidean distance between the closest points on the surfaces of the MRI/CT bone models and on the surface of the gold-standard model (i.e. deviations). For each knee, the deviation from the gold-standard was reported in terms of root mean square deviation (RMSD) and average deviation (AD), with a positive AD meaning that the MRI/CT model is on average larger than the gold-standard model.

A two-factor repeated measures analysis of variance (ANOVA) determined whether imaging modalities at three levels (MRI—2 mm, CT—0.625 mm, CT—1.25 mm) and registration method at two levels (ICP and fiducial markers) caused a change in the morphological errors (i.e. RMSD and AD) in 3D bone models of the femur and tibia. Tukey's test was used to compare the mean RMSD between the three levels of imaging modality. Statistical significance was considered when $p \leq 0.05$ for both the ANOVA and Tukey's tests.

Results

The imaging modality chosen to generate the 3D bone models caused significant changes in the morphological errors computed in terms of RMSD ($p < 0.0001$) and AD ($p < 0.0001$). The 3D femur and tibia models generated with CT differed from the corresponding GS model with an RMSD ranging from 0.3 to 0.5 mm, while the MRI models differed from the GS model with an RMSD ranging from about 1 to 1.2 mm (Tables 1 and 2, Figure 5). The CT models overestimated the size of the bone, with AD values ranging 0.1–0.2 mm, while the MRI models underestimated the size of the bone, with AD values ranging −0.7 to −0.8 mm (Tables 1 and 2, Figure 6). Interestingly, Tukey's test showed no differences between the 3D bone models generated using the two CT protocols for the RMSD ($p = 0.78$).

The registration method chosen to compare 3D bone models caused significant changes in the morphological errors in terms of RMSD ($p < 0.0001$) but not in terms of the AD ($p = 0.32$, Figure 6). On average the RMSDs found when registering the CT and MRI models using ICP were 0.1−–0.2 mm smaller than when using fiducial markers (Table 3).

Discussion

The morphological errors associated with 3D bone models of the femur and tibia generated from MRI and CT scans must be considered for applications based on these models. However, no previous studies have met the many methodological challenges necessary to accurately determine these errors. Hence,

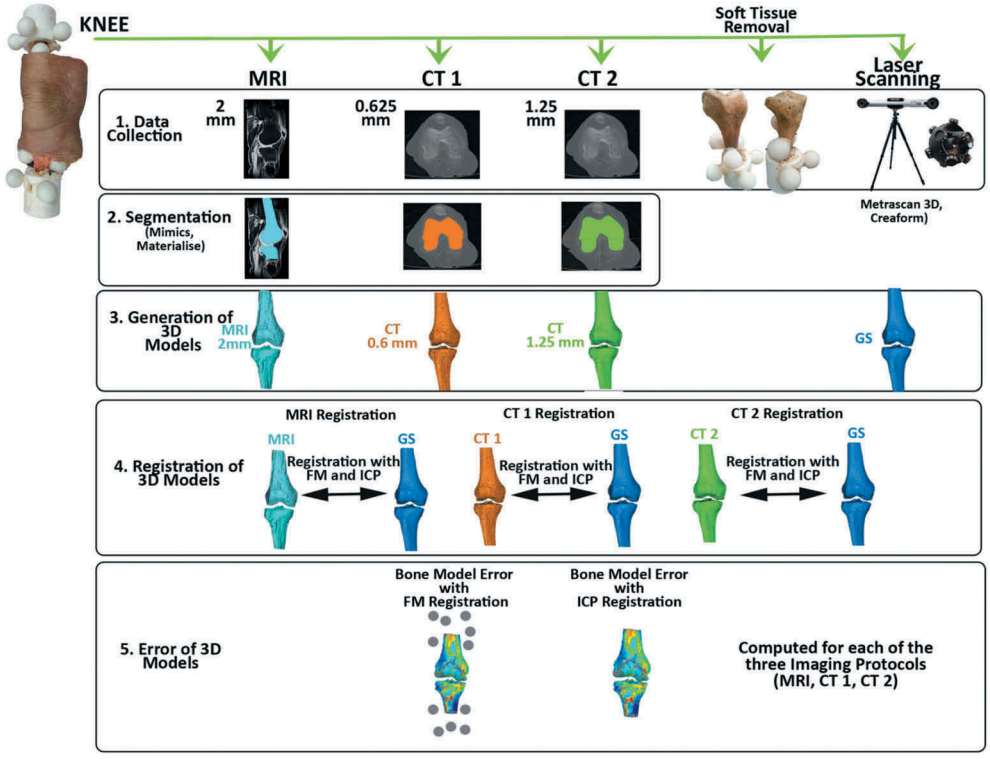

Figure 4. Flowchart of the data collection for the study. Deviations from the gold-standard model (GS) were evaluated for models generated using 1 MRI and 2 CT scans with two slice thicknesses, and for two methods of registration. One method used fiducial markers (FMs) and the other used the iterative closest point (ICP) algorithm.

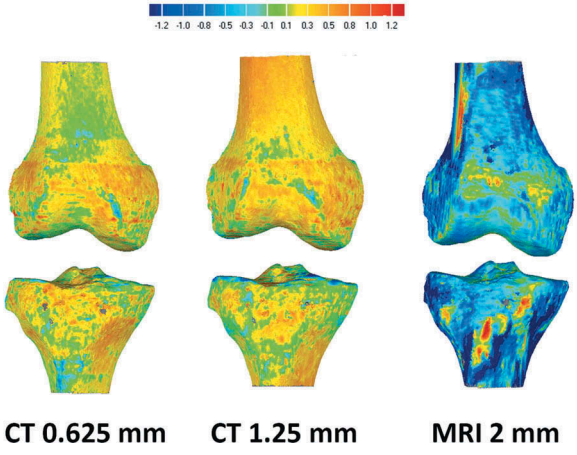

Figure 5. Deviation maps representing the morphological errors of the 3D bone models generated with CT and MRI for a representative knee specimen registered with the ICP algorithm. The comparison of the MRI model to the gold-standard (GS) model generated a blue-toned deviation map, meaning that the MRI bone model is smaller than the GS model. For the CT models, the comparison with the GS model generated a warm-toned deviation map, meaning that the CT bone model is slightly larger than the GS model.

the objectives of the present study were to determine the morphological errors of 3D bone models of the knee generated with clinically-adopted MRI and CT imaging protocols and to determine whether these morphological errors were different when using two 3D-3D registration methods and different CT slice thicknesses. The key findings were that 1) 3D bone models generated with CT using both 0.625 mm and 1.25 mm slice thicknesses yield submillimeter accuracy while 3D bone models generated with MRI yield accuracy worse than 1 mm; 2) CT bone models slightly overestimate the bone morphology, while the MRI bone models substantially underestimate it; 3) there is no difference in the morphological errors of 3D bone models generated with 0.625 mm and 1.25 mm slice thickness CT scans; and 4) ICP registration underestimates the morphological errors when compared to registration with fiducial markers.

This study has certain limitations. First, the segmentation method used in this study was more lengthy and detail-oriented than that typically used in a clinical application, in which segmentation generally cannot take more than 1 h. Accordingly our results represent a best-case scenario in terms of accuracy of the 3D bone models that can be generated with the images collected in this study. Second, the registration of the 3D models using fiducial markers was not error-free. Random errors that occurred in the scanning, segmentation and best-fitting spheres to the fiducial markers affected the correct identification of the centre of each fiducial marker, thus compromising the accuracy of the registration. Simulations to evaluate the effect of errors in the coordinates of the centre of the fiducial markers on the morphological errors (RMSD and AD) between two identical 3D femur models (Appendix A) revealed morphological errors of CT and MRI models of 0.07 mm and 0.23 mm for RMSD, respectively, and 0.0 mm for AD. Thus the true RMSD found for CT and MRI using fiducial markers as a registration method could be somewhat lower than the values reported in Tables 1 and 2 because a fraction of the morphological error found resulted from less than ideal registration of the fiducial markers.

Table 1. Morphological errors for the 3D femur models. The root mean square deviation (RMSD) and the average deviation (AD) are averaged across the 13 specimens and reported as mean ± standard deviation. The results for both the registration with fiducial markers (FMs) and the registration with the iterative closest point (ICP) algorithm are shown for each imaging protocol.

	CT 0.625 mm Femur	CT 1.25 mm Femur	MRI Femur
RMSD with FMs (mm)	0.5 ± 0.1	0.5 ± 0.1	1.2 ± 0.1
RMSD with ICP (mm)	0.3 ± 0.1	0.3 ± 0.1	1.0 ± 0.2
AD with FMs (mm)	0.2 ± 0.1	0.2 ± 0.1	−0.7 ± 0.2
AD with ICP (mm)	0.2 ± 0.1	0.2 ± 0.1	−0.8 ± 0.2

Table 2. Morphological errors for the 3D tibia models. The root mean square deviation (RMSD) and the average deviation (AD) are averaged across the 13 specimens and reported as mean ± standard deviation. The results for both the registration with fiducial markers (FMs) and the registration with the iterative closest point (ICP) are shown for each imaging protocol.

	CT 0.625 mm Tibia	CT 1.25 mm Tibia	MRI Tibia
RMSD with FMs (mm)	0.5 ± 0.1	0.4 ± 0.1	1.1 ± 0.2
RMSD with ICP (mm)	0.4 ± 0.1	0.3 ± 0.1	1.0 ± 0.2
AD with FMs (mm)	0.2 ± 0.1	0.1 ± 0.1	−0.7 ± 0.1
AD with ICP (mm)	0.2 ± 0.1	0.1 ± 0.1	−0.7 ± 0.2

This study indicates that when high-accuracy 3D bone models are needed, CT scanning is the better imaging modality to use with clinically adopted imaging protocols. CT bone models yielded submillimeter error because the average RMSD found was 0.5 mm. This error was most likely due to the limited image pixel size and slice thickness used, which were 0.39 mm and 0.625 mm, respectively. In contrast, MRI bone models yielded errors equal to or greater than 1 mm, so their use should be limited to applications whose outcome would not be markedly affected by such errors. The higher error found in the MRI bone models may be partly due to the larger pixel size of 0.78 mm and slice thickness of 2 mm, and to the physics of the MRI image generation that may be more prone to error than CT imaging.

Importantly, a systematic error occurred in the morphology of these models, with the CT and MRI bone models being 0.1–0.2 mm larger and 0.7–0.8 mm smaller, respectively, than the gold-standard. These errors may compromise the results of those applications that are sensitive to the systematic difference in morphology between the 3D bone model and that of the patient, such as patient-specific cutting guides. Although our study focused on morphological errors in bone models created from MRI, similar errors would be expected for bone-articular cartilage models created from MRI. Accordingly, a cutting or pinning guide whose design is based on an MRI bone-articular cartilage model may not seat on the patient's articular surface because it has been designed based on a bone-articular cartilage model which is generally smaller than that of the patient.

The fact that CT bone models generated with a 1.25 mm slice thickness yield comparable errors to those of CT bone models generated with a 0.625 mm slice thickness suggest that 1.25 mm slice thickness CT scans can be used for applications such as robotic-assisted total knee replacement (Pearle et al. 2009, Liow et al. 2014), patient-specific cutting guides (Hafez et al. 2006, Sassoon et al. 2015) and computational modelling (Donahue et al. 2002, Harrysson et al. 2007) without significant loss of accuracy but with the advantages of 1) reducing the radiation exposure to the patient and 2) having to store and segment only half of the images necessary for a 0.625 mm slice thickness CT scan. However one application where the larger slice thickness is not recommended is the generation of 3D bone models used for analysis of tibial contact kinematics in fluoroscopic studies (Komistek et al. 2003). In our experience, the use of the larger 1.25 mm slice thickness causes flattening of the distal articular surfaces of the femoral condyles thus introducing error in the computed contact locations.

The morphological errors of the 3D bone models were on average 20% systematically lower using the ICP algorithm than fiducial markers. Hence, when using the ICP algorithm to register two bone models, it is important to recognise that the ICP algorithm always reduces the error between the two bone models and most likely registers the two bone models in a relative position which is not the correct one if the morphology of the two 3D bone models differs. This is an expected result considering that the ICP algorithm is based on the minimisation of the morphological differences between two 3D models.

In comparing our results to those previously published, one previous study generated 3D bone models of nine human tibias using CT and MRI imaging protocols similar to those used in this study and compared them to gold-standard bone models generated with a laser scanner (Van Den Broeck et al. 2014). Their results for the proximal tibia showed a comparable error for the

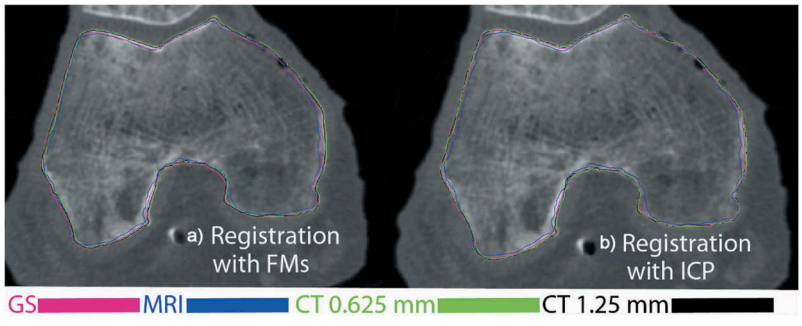

Figure 6. CT image (0.625 mm) of the femur with outlines of the 3D bone models of the gold-standard (GS), MRI, CT 0.625 mm and CT 1.25 mm. Figure 6(a) shows the outlines of the 3D bone models registered using fiducial markers (FMs), while Figure 6(b) shows the outlines of the 3D bone models registered using the iterative closest point (ICP). From these figures it can be noted that, for both registration methods, the MRI outline is on average smaller than the outline of the GS, while the CT 0.625 mm and CT 1.25 mm outlines are very close to each other and on average slightly larger than the outline of the GS.

Table 3. Differences between the morphological errors found using registration with fiducial markers (FMs) and registration with the iterative closest point (ICP) algorithm. Differences are averaged across the 13 specimens (FMs-ICP) are reported as mean ± standard deviation.

	CT 0.625 mm Femur	CT 0.625 mm Tibia	CT 1.25 mm Femur	CT 1.25 mm Tibia	MRI Femur	MRI Tibia
RMSD Difference (FMs-ICP, mm)	0.2 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.1 ± 0.1	0.2 ± 0.2	0.1 ± 0.2
AD Difference (FMs-ICP, mm)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.1	0.1 ± 0.1	0.1 ± 0.1

CT bone models to that found in this study (RMSD, 0.6 mm vs 0.5 mm), but a smaller error for the MRI bone models (RMSD, 0.6 mm vs 1.1 mm) (Van Den Broeck et al. 2014). In terms of AD, they found that both the CT and MRI bone models systematically overestimated the gold-standard bone model by 0.6 mm and 0.1 mm respectively, which is in contrast with our findings. However, the previous study used a maceration process which changed the size of the bone; the combined effect of the acetone treatment and boiling shrunk the bone specimen and thus the gold-standard bone model by an average −0.4 mm (Van Den Broeck et al. 2014). If we apply the −0.4 mm correction to the ADs found previously (Van Den Broeck et al. 2014), then the CT bone model would overestimate the gold standard model by 0.2 mm (0.6 mm–0.4 mm), and the MRI bone model would underestimate the gold standard model by 0.3 mm (0.1 mm–0.4 mm), which is in a closer agreement to our study. Another study found an AD equal to 0.6 mm for the CT model of the distal femur which is higher than in our study, but found that this inflated value was caused by the shrinking of the bone specimen due to boiling (Gelaude et al. 2008). Similar systematic errors with the CT and MRI 3D bone models overestimating and underestimating, respectively, the gold standard of the distal femur were reported (Rathnayaka et al. 2012).

In summary, this study demonstrated that CT-based bone models can be used for applications requiring submillimeter accuracy, while MRI-based bone models generated with the protocol used in this study should be used only when accuracy around 1 mm is acceptable. Further, because the errors in 3D models based on 0.625 mm and 1.25 mm slice thicknesses are comparable, the larger slice thickness can be used in certain applications without loss of accuracy. One exception is in the generation of 3D bone models to be used in fluoroscopic studies of tibial contact kinematics. Additionally, when the morphological errors must be determined for 3D bone models generated from protocols that are different from those used in this study, it is important to consider that using the ICP algorithm as a registration method will underestimate morphological errors.

Acknowledgments

The authors would like to thank individuals who donate their bodies for the advancement of education and research. This work was supported by the funding provided by THINK Surgical, Inc. which did not influence the study design, collection of data, interpretation of data and the decision for publication. Lastly, the authors would like to thank Rocio Lozano for her help in data collection and analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by THINK Surgical, Inc., Fremont, CA.

Notes on contributors

Valentina Campanelli, PhD completed her education at the University of Verona, Italy in 2015 and her postdoctoral scholarship at the University of California at Davis in 2016. Her research interests at this time in her career center on optimizing and validating surgical tools and procedures for orthopedic interventions, especially knee and hip replacements. She currently works as a clinical research manager at THINK Surgical Inc and has authored more than 10 peerreviewed papers in scientific journals.

Stephen M. Howell, MD completed his medical degree at Northwestern University in 1981 and his residency in orthopaedic surgery at Thomas Jefferson University in 1986. As a clinician, he specializes in treating knee injury and disease. As a researcher, he develops new surgical tools and techniques for treating knee injuries and disease particularly those to the soft tissues. He holds an appointment as an Adjunct Professor of Biomedical Engineering. He has published widely and has more than 125 publications in archival journals.

Maury L. Hull, PhD completed his education at the University of California Berkeley in 1976. He holds the position of Distinguished Professor and has appointments in three academic departments. His research interests at this time in his career center on developing improved surgical procedures for treating injuries and diseases to the soft tissues of the human knee. He is the author or co-author of more than 220 publications in scientific journals and is the recipient of a number of honors and awards including the H. R. Lissner Medal from the Bioengineering Division of the American Society of Mechanical Engineers. This is a career-level award for contributions to bioengineering research, education, and professional service.

References

Arun KS, Huang TS, Blostein SD. 1987. Least-squares fitting of two 3-D point sets. IEEE Trans Pattern Anal Mach Intell. 9:698–700.

Bellemans J, Banks S, Victor J, Vandenneucker H, Moemans A. 2002. Fluoroscopic analysis of the kinematics of deep flexion in total knee arthroplasty. J Bone Joint Surg Br. 2002:50–53.

Besl PJ, Mckay ND. 1992. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell. 14:239–256.

Campanelli V, Howell SM, Hull ML. 2016. Accuracy evaluation of a lower-cost and four higher-cost laser scanners. J Biomech. 49:127–131.

Cheng FB, Ji XF, Lai Y, Feng JC, Zheng WX, Sun YF, Fu YW, Li YQ. 2009. Three dimensional morphometry of the knee to design the total knee arthroplasty for Chinese population. Knee. 16:341–347.

Donahue TL, Hull ML, Rashid MM, Jacobs CR. 2002. A finite element model of the human knee joint for the study of tibio-femoral contact. J Biomech Eng. 124:273–280.

Fregly BJ, Rahman HA, Banks SA. 2005. Theoretical accuracy of model-based shape matching for measuring natural knee kinematics with single-plane fluoroscopy. J Biomech Eng. 127:692–699.

Gelaude F, Vander Sloten J, Lauwers B. 2008. Accuracy assessment of CT-based outer surface femur meshes. Comput Aided Surg. 13:188–199.

Gustke KA, Golladay GJ, Roche MW, Elson LC, Anderson CR. 2014. A new method for defining balance: promising short-term clinical outcomes of sensor-guided TKA. J Arthroplasty. 29:955–960.

Hafez M, Chelule K, Seedhom B, Sherman K. 2006. Computer-assisted total knee arthroplasty using patient-specific templating. Clin Orthop Relat Res. 444:184–192.

Harrysson OL, Hosni YA, Nayfeh JF. 2007. Custom-designed orthopedic implants evaluated using finite element analysis of patient-specific computed tomography data: femoral-component case study. BMC Musculoskeletal Disorders. 8:91.

ISO 5725-2. 1994. Accuracy (trueness and precision) of measurement methods and results part 2: basic method for the determination of repeatability and reproducibility of a standard measurement method.

Kang KT, Kim SH, Son J, Lee YH, Koh YG. 2017. Validation of a computational knee joint model using an alignment method for the knee laxity test and computed tomography. Biomed Mater Eng. 28:417–429.

Kaptein BL, Valstar ER, Stoel BC, Rozing PM, Reiber JH. 2003. A new model-based RSA method validated using CAD models and models from reversed engineering. J Biomech. 36:873–882.

Knee M. Medacta. https://media.medacta.com/media/newleafleftpazienti-au-99my2616au-rev2601.pdf.

Komistek RD, Dennis DA, Mahfouz M. 2003. In vivo fluoroscopic analysis of the normal human knee. Clinical Orthopaedics and Related Research. 410:69–81.

Kwak DS, Surendran S, Pengatteeri YH, Park SE, Choi KN, Gopinathan P, Han SH, Han CW. 2007. Morphometry of the proximal tibia to design the tibial component of total knee arthroplasty for the Korean population. Knee. 14:295–300.

Liow MH, Xia Z, Wong MK, Tay KJ, Yeo SJ, Chin PL. 2014. Robot-assisted total knee arthroplasty accurately restores the joint line and mechanical axis. A prospective randomised study. J Arthroplasty. 29:2373–2377.

Lorensen WE, Cline HE 1987. Marching cubes: A high resolution 3D surface construction algorithm. Proceedings of the ACM Siggraph '87 Computer Graphics; ACM.

Moro-Oka TA, Hamai S, Miura H, Shimoto T, Higaki H, Fregly BJ, Iwamoto Y, Banks SA. 2007. Can magnetic resonance imaging-derived bone models be used for accurate motion measurement with single-plane three-dimensional shape registration? J Orthop Res. 25:867–872.

Pearle AD, Kendoff D, Stueber V, Musahl V, Repicci JA. 2009. Perioperative management of unicompartmental knee arthroplasty using the MAKO robotic arm system (MAKOplasty). Am J Orthop. 38:16–19.

Rathnayaka K, Momot KI, Noser H, Volp A, Schuetz MA, Sahama T, Schmutz B. 2012. Quantification of the accuracy of MRI generated 3D models of long bones compared to CT generated 3D models. Med Eng Phys. 34:357–363.

Rathnayaka K, Sahama T, Schuetz MA, Schmutz B. 2011. Effects of CT image segmentation methods on the accuracy of long bone 3D reconstructions. Med Eng Phys. 33:226–233.

Riley J, Roth JD, Howell SM, Hull ML. 2018a. Increases in tibial force imbalance but not changes in tibiofemoral laxities are caused by varus-valgus malalignment of the femoral component in kinematically aligned TKA. Knee Surg Sports Traumatol Arthrosc. doi:10.1007/s00167-018-4841-6

Riley J, Roth JD, Howell SM, Hull ML. 2018b. Internal-external malalignment of the femoral component in kinematically aligned total knee arthroplasty increases tibial force imbalance but does not change laxities of the tibiofemoral joint. Knee Surg Sports Traumatol Arthrosc. 26:1618–1628.

Sassoon A, Nam D, Nunley R, Barrack R. 2015. Systematic review of patient-specific instrumentation in total knee arthroplasty: new but not improved. Clin Orthop Relat Res. 473:151–158.

Steadman DW, DiAntonio LL, Wilson JJ, Sheridan KE, Tammariello SP. 2006. The effects of chemical and heat maceration techniques on the recovery of nuclear and mitochondrial DNA from bone. J Forensic Sci. 51:11–17.

Truematch Personalized Solutions. Depuy synthes. synthes.vo.llnwd.net/o16/LLNWMB18/.../0612-0617-0510r0613_TM%0620RG%0620DR.pdf.

Van Den Broeck J, Vereecke E, Wirix-Speetjens R, Vander Sloten J. 2014. Segmentation accuracy of long bones. Med Eng Phys. 36:949–953.

White D, Chelule KL, Seedhom BB. 2008. Accuracy of MRI vs CT imaging with particular reference to patient specific templates for total knee replacement surgery. Int J Med Robot. 4:224–231.

Appendix

A. Maceration and Segmentation Repeatability

To ensure that no change in the bone morphology occurred during the cleaning process via dermestid beetles, a CT scan of three specimens was also performed after maceration. The same CT protocol with 0.625 mm slice thickness, and the same segmentation protocol used for the CT before soft tissue removal were used for this scan. The 3D bone models generated from the CT scans performed before and after maceration were registered using fiducial markers as described above, and differences in morphology were determined in terms of RMSD and AD. Furthermore, the repeatability error in the segmentation of the CT images was computed to ensure that the differences found in the CT models before and after maceration were not due to repeatability error in the segmentation. The same set of CT images was segmented five times and the resulting five 3D models were compared using a pairwise comparison (i.e. ten comparisons) so that the average RMSD and AD values could be computed.

Small differences were found between the morphology of the CT bone models generated before and after the soft tissue removal via dermestid beetles (Table A1), with an average RMSD of 0.3 mm and an average AD of 0.1 mm. These differences are within the repeatability error generated during segmentation, which were 0.3 mm for RMSD and 0.1 mm for AD (Table A1). Hence, the maceration using dermestid beetles did not introduce any quantifiable change in morphology.

Table A1. Morphological differences between 3D models generated with CT before and after soft tissue removal with dermestid beetles (CT femur dermestid beetles and CT tibia dermestid beetles) and morphological differences between 3D models developed from identical sets of CT images (CT femur segmentation repeatability and CT tibia segmentation repeatability). The root mean square deviation (RMSD) and the average deviation (AD) averaged across the three specimens and across the ten pairwise comparisons, respectively, are reported as mean ± standard deviation.

	RMSD (mm)	AD (mm)
CT femur dermestid beetles	0.3 ± 0.1	0.1 ± 0.1
CT tibia dermestid beetles	0.3 ± 0.0	0.1 ± 0.0
CT femur segmentation repeatability	0.3 ± 0.0	0.1 ± 0.1
CT tibia segmentation repeatability	0.3 ± 0.1	0.1 ± 0.1

B. Simulation of Registration Error with Fiducial Markers

Because 3D-3D registration using fiducial markers is not an error-free process, simulations were performed to evaluate the effect of errors in the coordinates of the fiducial marker centres on the morphological errors (RMSD and AD) found between two identical 3D bone models. Error in the fiducial marker centres mostly derives from random noise intrinsic to the CT/MRI scanners or laser scanners, the variability of the segmentation performed in Mimics (for CT/MRI models only), and the variability of the selection of the surface of the 3D marker model chosen to best fit a sphere (i.e. some regions of the 3D model of a fiducial marker are not spherical and should not be selected for the best fit).

The simulations entailed several steps. The first step was to generate the coordinates of the centres of a set of seven fiducial markers and add noise to each of the seven sets of coordinates to simulate the random error generated during the laser scanning of the markers and the best-fitting of the spheres. The second step was to add a different amount of noise to a duplicate of the original set of seven marker centres to simulate the noise generated during MRI/CT scanning, the segmentation of the markers, and the best-fitting of the spheres. Each error was randomly sampled from a normal distribution with a mean of 0 mm and a standard deviation of 0.04 mm for the laser scanned set, and standard deviations of 0.05 mm and 0.25 mm for the CT and MRI sets, respectively. These noise levels were chosen based on a test in which the same specimens were CT-scanned, MRI-scanned, and laser-scanned five times each. For each scan, the location of each fiducial marker centre was identified and the repeatability error of the location of each fiducial marker centre for each imaging modality was computed according to ISO standard 5725-2 (ISO 5725-2 1994: 1994). The third step in the simulation required registering the two sets of fiducial markers (the laser-scanned set with the CT set, as well as the laser-scanned set with the MRI set) using a least-square fitting algorithm (Arun et al. 1987) to generate a transformation matrix to register two identical 3D bone models, and compute the RMSD and AD between the two models. Ideally, because the two 3D models used in the simulation are identical, the AD and RMSD should all be zero. Two hundred simulations were performed for each of the CT and MRI femurs.

The simulations showed that for CT, the error in the coordinates of the fiducial marker centres introduces an increase in the morphological error of 0.07 mm for RMSD and 0.0 mm for AD. The simulations showed that for MRI, the error in the coordinates of the fiducial marker centres introduces an increase in the morphological error of 0.23 mm for RMSD and 0.0 mm for the AD. For both CT and MRI, these results may explain in part the 0.1–0.2 mm difference in RMSD found between the two registration methods (Table 3). Also the true RMSD found for CT and MRI could actually be somewhat lower than the values reported in Tables 1 and 2 because a fraction of the morphological error found was generated as a consequence of the random error in the coordinates of the fiducial marker centres.