This is the author's version of a work that was submitted/accepted for publication in the following source:


This file was downloaded from: http://eprints.qut.edu.au/45764/

© Copyright 2012 Elsevier Ltd

This is the author's version of a work that was accepted for publication in Medical Engineering and Physics. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Medical Engineering and Physics, Medical Engineering and Physics, VOL34, ISSUE3, (2012) DOI: 10.1016/j.medengphy.2011.07.027.

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source:

http://dx.doi.org/10.1016/j.medengphy.2011.07.027
Quantification of the accuracy of MRI generated 3D models of long bones compared to CT generated 3D models

Kanchana Rathnayaka¹, Konstantin I. Momot², Hansrudi Noser³, Andrew Volp⁴, Michael A. Schuetz¹, Tony Sahama² and Beat Schmutz¹

1. Institute of Health and Biomedical Innovation, 60 Musk Avenue, Kelvin Grove, QLD 4059, Australia
2. Queensland University of Technology, 2 George Street, Brisbane, QLD 4000, Australia
3. AO Research Institute Davos, Clavadelerstrasse 8, 7270 Davos, Switzerland
4. Princess Alexandra Hospital, 199 Ipswich Road, Woolloongabba, Brisbane, QLD 4102, Australia

Corresponding author:
Beat Schmutz
Institute of Health and Biomedical Innovation
60 Musk Avenue, Kelvin Grove
QLD 4059, Australia
Phone: +61 7 3138 6238
Fax: +61 7 3138 6030
Email: b.schmutz@qut.edu.au
Abstract
Orthopaedic fracture fixation implants are increasingly being designed using accurate 3D models of long bones based on computer tomography (CT). Unlike CT, magnetic resonance imaging (MRI) does not involve ionising radiation and would therefore be a desirable alternative to CT. This study aims to quantify the accuracy of MRI-based 3D models compared to CT-based 3D models of long bones. The femora of five intact cadaver ovine limbs were scanned using a 1.5T MRI and a CT scanner. Image segmentation of CT and MRI data was performed using a multi-threshold segmentation method. Reference models were generated by digitising the bone surfaces free of soft tissue with a mechanical contact scanner. The MRI- and CT-derived models were validated against the reference models. The results demonstrated that the CT-based models contained an average error of 0.15 mm while the MRI-based models contained an average error of 0.23 mm. Statistical validation shows that there are no significant differences between 3D models based on CT and MRI data. These results indicate that the geometric accuracy of MRI based 3D models was comparable to that of CT-based models and therefore MRI is a potential alternative to CT for generation of 3D models with high geometric accuracy.

Keywords: MRI, CT, 3D Models, Femur
1. Introduction

Three-dimensional (3D) models of long bones with a high geometric accuracy are widely utilised by medical engineering research and in clinical practice. The design of orthopaedic fracture fixation implants [1, 2], computer aided surgery simulations [3, 4] and fracture healing models [5, 6] are just a few examples. Computed Tomography (CT) has become the gold standard for scanning of bones to produce 3D models with high geometric accuracy. Due to high radiation exposure, CT cannot be used to scan healthy human volunteers. Therefore, an alternative method for the scanning of long bones of the healthy human population needs to be investigated.

Among various uses of 3D models, orthopaedic implant design particularly requires 3D models with high geometric accuracy to produce implants with a better fit to the patients’ anatomy [1, 3]. Furthermore, the anatomically pre-shaped implants are often designed based on Caucasian population and thus the size and shape do not accurately match the Asian population. Therefore, those pre-shaped implants still need some optimisation for a better anatomical fit to people of different ethnic origins and age groups [2]. The ethnicity and age are two important factors which determine the shape and size of bones [7, 8]. Thus, a database with accurate bone data from different ethnic and age groups is essential for this purpose.

Some institutions have already started developing such data bases using CT imaging of cadaver bones [9] but the majority of these bones are from older donors (>60 years) therefore do not represent the young patient population. Furthermore, cadaver bones can seldom be chosen according to the researchers’ need (e.g. gender or specific subject height) due to the limited availability. Therefore, there is a need to collect bone data from healthy human volunteers who represent that part of the patient population for which no CT data exists or can be acquired. This would facilitate researchers’ access to specific population groups for the purpose of obtaining high-quality anatomical image data.

CT scanning of healthy human volunteers is not ethically justifiable due to the high radiation exposure [10, 11]. Studies investigating CT imaging protocols that use low radiation doses, while keeping the original image quality, have become an important part in modern research; however, the radiation exposure cannot be eliminated completely [12, 13]. Moreover, some countries do not allow scanning of healthy volunteers using any form of radiation (low or high dose) unless otherwise required for medical reasons. MRI has been used for scanning of soft tissues with high image quality for the purpose of generating anatomical 3D models [14, 15]. MRI is ideally suited for the scanning of soft tissues using proton (\(^1\)H) nuclei as the source of the signal. MRI is not routinely used for the scanning of bones, as the bone tissue does not generate a MR signal due to extremely short \(^1\)H transverse relaxation times (\(T_2\)). However, using the signal generated from the surrounding soft tissue, the geometry of the bone cortex can be identified (Figure 1). The main advantage of using MRI is that it
does not involve ionising radiation and therefore it is ideally suited for the scanning of healthy human volunteers for research purpose.

Figure 1. CT (left) and MRI (right) axial slices (the proximal end of the tibia is also visible on the right side of the images as the limb has been scanned with flexed knee joint) from the distal region of the femur. Cortical bone does not generate an MRI signal and appears in black while this is clearly visible in the CT image. Parts of the trabecular structure are visible due to the presence of bone marrow.

The following recent studies have demonstrated the potential of MRI for indirect bone imaging. In a study conducted by Moro-Oka et al, MRI has been used in conjunction with CT to generate 3D models of the knee joint to validate 3D kinematic measurements of single plane radiography [16]. The models included the distal part of the femur and proximal part of the tibia of three human volunteers. The geometric differences between MRI and CT outer 3D models were 0.11 ± 0.81 mm, 0.23 ± 0.48 mm and 0.12 ± 0.60 mm for the femora and 0.14 ± 0.67 mm, 0.13 ± 0.48 mm and 0.15 ± 0.77 mm for the tibiae. Musculoskeletal models which represent bone, cartilage, muscles and ligaments have been successfully generated using MRI and CT [17, 18] in which surface matching accuracy of MRI models to CT was 0.7 ± 0.1 mm. MRI has also been used for computer assisted spinal surgeries [19], foot bone motion quantification [20] and has provided an accuracy sufficient for the intended applications. However, these studies have not used MRI alone to generate the complete long bone as well as they have not been validated using an accurate reference standard such as a contact scanner or laser surface scanner. Therefore, further investigation using long bones is necessary before using MRI as a successful imaging technique for long bone imaging.

The aim of this study was comprehensive quantification of the accuracy of 3D MRI models of outer and inner surfaces of long bones by utilising state-of-the-art dense surface scans of the bones as the reference standard. The accuracy of different anatomical regions of the 3D models was quantified in
order to identify the regions of the bone that can be segmented accurately using MRI. We discuss the requirements for accurate MR anatomical imaging of long bones; we also discuss the principal challenges of this task, particularly the difficulty of accurate delineation of bone geometry when multiple tissues are present near the bone – soft tissue interface.

2. Methods

Five intact cadaver ovine hind limbs with intact soft tissues were used for this study, however, only the femora were imaged with CT and MRI. The limbs were amputated with part of the pelvis attached to ensure that the femoral head is not exposed to air. The sheep were 6-8 years old and had a mean weight of 34 kg. The average length of a femur was 17.6 cm and the average diameter was 1.8 cm. The specimens were kept frozen between procedures and each specimen was allowed to defrost at room temperature for 24 hrs before the scanning and dissection procedures.

MR images of the femora were obtained using a 1.5T clinical MRI scanner (Siemens Magnetom Avanto) with a body matrix RF coil using a 3D FLASH sequence, TR = 11 ms, TE = 4.94 ms, flip angle = 15°, slice thickness = 1 mm, pixel size = 0.45 mm × 0.45 mm. CT scanning (Phillips, Brilliance 64) of the femora were conducted using kVp = 140, mAs = 231, pixel size of 0.4 × 0.4 mm, slice spacing of 0.5 mm and a sharp convolution kernel (kernel D for Phillips scanners). In both scanning procedures, the long axis of the femur was visually aligned with the long axis of the CT scanner or the static magnetic field of the MRI magnet using scout views. Both MRI and CT data were saved in DICOM image format for further processing.

The segmentation of the CT and MRI image data was performed using a multi-threshold segmentation method [21]. This method calculates the average intensity level of the pixels which make up the edge detected by the Canny edge detection filter [21, 22]. In this study, the Canny filter was run on three randomly selected axial slices from each of (three) anatomical regions (Figure 2) for which the threshold values need to be calculated. Using the XY coordinates of the pixels which make up the edge, the intensity values of the corresponding pixels were read and the average of them was used as the threshold level for the particular region.

Figure 2. Three regions of thresholding

Using the three threshold levels calculated for each of the anatomical regions, segmentation of the MRI data was carried out in Amira 5.2 (Visage image) softwares system (Figure 3). The same multi-
threshold technique was used for the segmentation of the CT images using three different threshold levels calculated for CT data. Triangular meshed surfaces were generated from the segmented data and exported as STL files for further processing. Complete 3D models were reconstructed for the outer contour while only the diaphysis was reconstructed for the inner contour. Out of five bones, three were not used for generating the inner contour due to inadequate contrast on MR images between medullary canal and the cortical bone, as the bone-marrow did not generate a signal strong enough to differentiate between bone and bone-marrow.

Figure 3. Three orthogonal views (axial, coronal and sagittal) of thresholded bone from CT (above) and MRI (below) (only the proximal half of the bone is shown in coronal and sagittal views)

Table 1: Threshold values used for multi-thresholding of MRI Data.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Threshold Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proximal Region</td>
</tr>
<tr>
<td>Bone 1</td>
<td>92</td>
</tr>
<tr>
<td>Bone 2</td>
<td>93</td>
</tr>
<tr>
<td>Bone 3</td>
<td>110</td>
</tr>
<tr>
<td>Bone 4</td>
<td>143</td>
</tr>
<tr>
<td>Bone 5</td>
<td>147</td>
</tr>
</tbody>
</table>
A reference standard was generated using a mechanical contact scanner for the validation of the 3D models generated from the CT and MRI data. After CT and MRI scanning, the limb was dissected to obtain the bone free of soft tissue while care was been taken to preserve the bone surface. The denuded femora were scanned with a mechanical contact 3D scanner (Roland MDX 20) using a resolution of 0.3 mm × 0.3 mm in the scanning (X,Y) plane and 0.025 mm step size in which the stylus perpendicular to the X,Y plane approaches the bone surface. The bone was scanned in several stages generating a number of surfaces depending on the complexity of the bone geometry. These surfaces were then aligned with each other and reconstructed into a final 3D model using the reverse engineering software Rapidform 2006 (INUS Technology, Korea).

To validate the 3D models of the medullary canal, microCT (µCT-40, Scanco Medical, Switzerland)-generated models were used as the reference standard. The diaphysis of the denuded femur was scanned using an isotropic voxel size of 0.03 mm³, kVp = 70. Segmentation of the microCT data was carried out using the Canny edge detection filter, as a highly repeatable segmentation method, for delineating the inner contour of the diaphysis [21]. To remove the salt and pepper noise associated with microCT images, a 20×20 median filter was applied to each DICOM image using Matlab software package.

Before comparison, the 3D models were aligned to the reference 3D models using functions built in Rapidform 2006. First, the 3D models were brought into gross alignment moving the models with the track ball function. Then the fine registration function, which is based on the iterative closest point (ICP) algorithm [17, 23], was applied for the final proper alignment.

Geometric differences between the model of interest and the reference model were calculated using a point to point comparison method implemented in Rapidform 2006. The absolute (unsigned) geometric differences were recorded as the average distance with the standard deviation. The outer 3D models generated from the MRI and CT data were validated against the reference models generated from the mechanical contact scanner. The MRI-based models also were compared against the CT-based models to quantify the difference between the 3D models generated from the two imaging techniques. The inner 3D models generated from MRI and CT data were validated against the microCT generated inner 3D models using the microCT models as the reference standard. Comparison of the 3D models of the outer cortex were conducted as whole models as well as in five different anatomical regions (Figure 4) to localise the errors associated with different anatomical regions.
Figure 4: Five different anatomical regions of the femur: 1- head, 2- proximal region, 3- diaphysis, 4- distal region, 5- distal articular region.

Statistical differences between accuracies of 3D models based on MRI and CT were calculated using one way ANOVA. We hypothesised that there is no significant difference between the average deviations of CT vs reference and MRI vs reference models. The level of statistical significance was set to \( p \leq 0.05 \). The statistical validation was conducted between whole bone models as well as between different anatomical regions of the bone.

3. Results

Comparison of the CT-based 3D models (Figure 5) generated using the multi-threshold method with the reference models (contact scanner based models) showed that the two sets of models exhibited an average error of 0.15 mm (Figure 6). In comparison, the MRI-based 3D models (Figure 6) presented an average error of 0.23 mm when compared to the reference models. Comparison of the MRI-based models to the CT-based models generated an average error of 0.23 mm. Statistically there were no significant difference between the average deviations of CT and MRI models (\( p = 0.067 \)).

Figure 5: The final 3D model generated from MRI data (red) aligned with the reference model (green) on left and comparison of the surface geometry of those two models on right.
Figure 6: The average deviation between CT and reference, MRI and reference, and MRI- and CT-based 3D models.

When different anatomical regions of the CT-based and MRI-based 3D models were compared with the corresponding reference models, the diaphyseal region presented the smallest average errors of 0.07 mm and 0.15 mm respectively (Figure 7).

Figure 7: Comparison of different anatomical regions of the CT- and MRI-based models with the reference models.

According to the results, 50.9% surface area of CT models and 78.8% surface area of MRI models underestimated the reference model while 75.8% of the surface area of MRI models underestimated the CT models (Table 2). The proximal, diaphyseal and distal regions of CT models overestimated the reference models while the same regions of MRI models underestimated the reference models.
Articular regions (head and distal articular region) of both MRI- and CT-based models underestimated those regions of the reference model by similar percentages (Table 2).

Table 2: Percentage surface area of CT and MRI models that under- or over-estimated the reference model and percentage surface area of MRI model that under- or over-estimated the CT model.

<table>
<thead>
<tr>
<th>Region</th>
<th>CT vs Reference</th>
<th>MRI vs Reference</th>
<th>MRI vs CT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole bone</td>
<td>-50.9%</td>
<td>-78.8%</td>
<td>-75.8%</td>
</tr>
<tr>
<td>Head</td>
<td>-80.9%</td>
<td>-81.8%</td>
<td>-62.2%</td>
</tr>
<tr>
<td>Proximal region</td>
<td>66.8%</td>
<td>-70.5%</td>
<td>-79.7%</td>
</tr>
<tr>
<td>Diaphysis</td>
<td>60.6%</td>
<td>-93.8%</td>
<td>-95.2%</td>
</tr>
<tr>
<td>Distal region</td>
<td>58.2%</td>
<td>-68.8%</td>
<td>-71.6%</td>
</tr>
<tr>
<td>Distal articular region</td>
<td>-84.4%</td>
<td>-75.2%</td>
<td>-54.9%</td>
</tr>
</tbody>
</table>

Negative values indicate that the model of interest underestimates the reference model.

Statistically the differences were significant between distal regions (p = 0.008) of the CT and MRI models based on multi-threshold method (Table 3).

Table 3: Statistical significance between average deviations of CT vs reference and MRI vs reference models.

<table>
<thead>
<tr>
<th>Region</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>p = 0.065</td>
</tr>
<tr>
<td>Proximal region</td>
<td>p = 0.478</td>
</tr>
<tr>
<td>Diaphysis</td>
<td>p = 0.435</td>
</tr>
<tr>
<td>Distal region</td>
<td>p = 0.008</td>
</tr>
<tr>
<td>Distal articular region</td>
<td>p = 0.697</td>
</tr>
</tbody>
</table>

When the CT-based inner models were compared with the micro CT generated inner models, they presented an error of 0.10 mm (Figure 8) while the MRI-based models presented an error of 0.08 mm.
Figure 8: Comparison of the MRI- and CT-based inner models with the microCT models and MRI models with the CT models.

Figure 9. A cross section of the MRI (red) and reference (green) models through the diaphysis.

4. Discussion
Orthopaedic implant design and other medical engineering applications typically use 3D bone models with a high geometric accuracy that are generated from CT scan data, the current gold standard of acquiring data for 3D modelling. However, for ethical reasons, CT cannot be used to scan healthy volunteers since CT exposes the individual to a high dose of ionising radiation. Therefore, CT can
only be used on cadaver bones which mostly represent the elderly (>60 yrs) population. One feasible alternative for this is to use MRI for acquisition of data from healthy individuals who can then be selected according to study specific protocols. Therefore, this study aimed to quantitatively validate the accuracy of MRI generated 3D models of long bones using ovine femora.

Using five ovine femora, this study demonstrated a comparable accuracy between 3D models generated from MRI and CT data for long bones. MRI models generated using a multi-threshold segmentation method presented a mean deviation of 0.23 mm when compared to the reference model (Figure 6). In comparison, CT-based model presented a mean deviation of 0.15 mm when the same thresholding method is used. Even though the mean deviation of MRI generated models are slightly higher than that of CT-based models, the differences were not statistically significant ($p = 0.067$). In general, about 78% of surface area of the MRI-based models and 51% of surface area of the CT-based models underestimated the reference models (Figure 9). The average deviation of $0.23 \pm 0.04$ mm between MRI and CT models was smaller to the average deviation of $0.7 \pm 0.1$ mm obtained for a similar study conducted by Lee et al [17]. This difference might be the result of manual segmentation used by Lee et al for the generation of their MRI-based 3D models.

The inside models also generated comparable results to CT, 0.08 mm and 0.10mm for MRI and CT models respectively when compared with microCT-based inside models. One drawback of the inside model comparison was that the limited number of samples used due to the insufficient contrast of three of the samples. Even though these three samples suffered from inadequate contrast, in our experience, scanning of human long bones with MRI clearly produced a higher contrast between medullary canal and cortical bone [24]. When the errors associated with different anatomical regions were quantified, the diaphysis of both MRI- and CT-based models presented the highest accuracy, while the head and distal articular regions showed the lowest accuracy. Except for the distal region, there were not any significant statistical differences between all the other regions of the MRI and CT models.

In order to understand the differences between the imaging accuracy in different anatomical regions of the femur, one must consider the nature of the bone-soft tissue interface. In the diaphyseal region, the bone boundary is mostly a simple two-tissue interface between the bone and the muscle. On the other hand, in the proximal and distal regions, where the bone forms part of a movable joint, multiple tissues are present (e.g. muscle, fat, cartilage, synovial fluid, tendons and ligaments). Some of these tissues have different visibility in MR and CT images: for example, articular cartilage is invisible in CT and may or may not be visible in MRI, depending on the technique used. The partial volume effect (where a single voxel contains more than one tissue type) blurs the interface between tissues; the resulting loss of information is especially severe in regions where multiple tissues (and multiple
interfaces) meet. Therefore, a definition of the “boundary of the bone” in imaging of the proximal and distal regions is inherently based on incomplete information and often requires interpolation of the “missing” fragments of interfaces between tissues.

Owing to this, the segmented MRI images in this work had to be manually corrected by removing the additionally selected structures (e.g. tendons). This manual editing is a potential source of error in the reconstructed 3D models. The task of delineating the bone geometry is further complicated by the fact that the $T_2$ of articular cartilage is orientation-dependent [25, 26], which may impart different appearances to parts of the same cartilage that are oriented at different angles to the static magnetic field. The articular cartilages however were not included either in CT or MRI image segmentation as the cartilage is not visible in CT images while the cartilages were included in the reference model. This explains the higher inaccuracies of the articular regions of the 3D models. The femoral head and the distal articular regions underestimated the reference model for \(-80\%\) of the surface area, which also confirms that the cartilage has not been included in CT or MRI models while they were being segmented.

Another important factor that affects the accuracy of the 3D models is the uniformity of the imaging gradients of the MRI spectrometer. The region where the gradient is uniform is limited; this imposes an upper limit on the field of view compatible with accurate reconstruction of the geometry. Since the effective scanning length of samples is limited to 25-30 cm, a human long bone has to be scanned in several stages depending on the length of the sample. In addition, the larger slice spacing of the MRI data (1 mm) compared to that of CT data (0.5mm) influences the accuracy of the models reconstructed from more complex geometric shapes (e.g. femoral head, distal articular region) than the simple geometric shapes (bone shaft) [27, 28]. In the long term, the development of highly linear gradients and precise slice selection techniques is desirable for anatomical imaging of long bones because these developments would significantly increase the efficiency of MR bone imaging.

Although the accuracy is comparable to CT and there are no radiation hazards, MRI requires longer scanning times compared to CT. For example; MRI scanning of an ovine femur usually requires about 20 minutes of scanning time whereas CT scanning of the same sample takes only a few seconds. Therefore, MRI scanning of human volunteers is vulnerable to artefacts caused by the volunteer moving the leg which might have an effect on the accuracy of the 3D models. This effect has already been observed by our research group and further studies are being conducted to correct these artefacts.

This study utilised ovine femora as samples that are \(~17.6\) cm in length and \(~1.8\) cm in diameter. Even though the ovine femora (17.6 cm) are considerably shorter than human femora (46.9 cm) [29] or humeri (29.8 cm) [30], the diameter of the ovine femora (1.8 cm) is similar to the diameter of the human humeri (1.8 cm) [30] though smaller to the diameter of human femora (2.9 cm antero-posterior
and 2.8 cm medio-lateral) [31]. Therefore, it is reasonable to expect that the methods used in this study can be applied to generate 3D models from MRI of human long bones with similar accuracies.

This study aimed to quantify the accuracy of the MRI generated 3D models of long bones compared to the CT-based models using five ovine femora with reference to localisation of the errors associated with different anatomical regions of the bone. The results revealed that the accuracy of the MRI derived long bone models is comparable to CT-derived models. This indicates that MRI is a potential alternative to generate 3D models from long bones with a comparable geometric accuracy to CT. However, further investigations are necessary to overcome the drawbacks associated with the long scanning times.

5. Acknowledgement
This study was supported in part through funding from the AO Research Institute. The last author has received an industrial scholarship from Synthes Asia-Pacific.

6. References


