HHS Public Access

Author manuscript

Bull Hosp Jt Dis (2013). Author manuscript; available in PMC 2020 July 06.

Published in final edited form as:

Bull Hosp Jt Dis (2013). 2019 March ; 77(2): 115–121.

A Novel MRI Tool for Evaluating Cortical Bone Thickness of the Proximal Femur

Austin J. Ramme, MD, PhD,
Department of Orthopedic Surgery, NYU Langone Orthopedic Hospital, New York, New York, USA

Shaleen Vira, MD,
Department of Orthopedic Surgery, NYU Langone Orthopedic Hospital, New York, New York, USA

Alexandra Hotca, BA,
Department of Radiology, Center for Biomedical Imaging, NYU Langone Medical Center, New York, New York, USA

Rhiannon Miller,
Department of Radiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA.

Arakua Welbeck, BS,
Department of Radiology, Center for Biomedical Imaging, NYU Langone Medical Center, New York, New York, USA

Stephen Honig, MD,
Osteoporosis Center, NYU Langone Orthopedic Hospital, New York, New York, USA

Kenneth A. Egol, MD,
Department of Orthopedic Surgery, NYU Langone Orthopedic Hospital, New York, New York, USA

Chamith S. Rajapakse, PhD,
Department of Radiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA.

Gregory Chang, MD
Department of Radiology, Center for Biomedical Imaging, NYU Langone Medical Center, New York, New York, USA

Abstract

Correspondence: Gregory Chang, MD, NYU Langone Medical Center, 660 First Ave, 3rd Floor, Room 333, New York, New York 10016, USA; gregory.chang@nyumc.org.

Disclosure Statement

None of the authors have a financial or proprietary interest in the subject matter or materials discussed herein, including, but not limited to, employment, consultancies, stock ownership, honoraria, and paid expert testimony.

Background—Osteoporotic hip fractures heavily cost the health care system. Clinicians and patients can benefit from improved tools to assess bone health. Herein, we aim to develop a three-dimensional magnetic resonance imaging (MRI) method to assess cortical bone thickness and assess the ability of the method to detect regional changes in the proximal femur.

Methods—Eighty-nine patients underwent hip magnetic resonance imaging. FireVoxel and 3DSlicer were used to generate three-dimensional proximal femur models. ParaView was used to define five regions: head, neck, greater trochanter, intertrochanteric region, and subtrochanteric region. Custom software was used to calculate the cortical bone thickness and generate a color map of the proximal femur. Mean cortical thickness values for each region were calculated. Statistical t-tests were performed to evaluate differences in cortical thickness based on proximal femur region. Measurement reliability was evaluated using coefficient of variation, intraclass correlation coefficients, and overlap metrics.

Results—Three-dimensional regional cortical thickness maps for all subjects were generated. The subtrochanteric region was found to have the thickest cortical bone and the femoral head had the thinnest cortical bone. There were statistically significant differences between regions ($p < 0.01$) for all possible comparisons.

Conclusions—Cortical bone is an important contributor to bone strength, and its thinning results in increased hip fracture risk. We describe the development and measurement reproducibility of an MRI tool permitting assessment of proximal femur cortical thickness. This study represents an important step toward longitudinal clinical trials interested in monitoring the effectiveness of drug therapy on proximal femur cortical thickness.

Cortical bone is an important contributor to bone strength, and thinning of cortical bone is an established risk factor for hip fracture.[1] Computed tomography (CT) has traditionally been considered the standard for bone imaging due to the high contrast between bone and surrounding soft tissues. This is useful to create virtual three-dimensional surface models of bone[2] that have been used for patient-specific finite element analysis,[3] implant design,[4] and volumetric characterization of disease states.[5,6] However, CT scanners administer ionizing radiation and because the administration of ionizing radiation must be restricted, there are limits to the imaging resolution that can be attained with CT in vivo as well as limits to the number of CT scans that could be serially performed if one desires to monitor disease progression or therapy response.

Magnetic resonance imaging (MRI), which detects as an electrical signal the changing magnetic field generated within a tissue sample, does not expose patients to ionizing radiation. At conventional echo times, images of bone can be generated via the MRI signal detected from marrow (fat and water) and surrounding soft tissues, which provides contrast to bone tissue, which itself is hypointense or dark. High-resolution imaging of bone at the microarchitectural level has become possible due to improvements in MRI radiofrequency coil technology that permits increased signal-to-noise ratio. This has been described in the distal radius and tibia[7–9] and more recently in the hip.[10,11] The hip is one of the most clinically important anatomic sites in musculoskeletal disease as osteoporotic hip fracture and hip osteoarthritis procedures cost the U.S. health care system tens of billions of dollars annually.[12,13]

The ability to develop three-dimensional cortical bone thickness maps has allowed for the evaluation of implants and anatomic changes seen in different musculoskeletal diseases. Tucker et al.[14] used three-dimensional CT imaging to investigate changes in femoral shaft cortical thickness for intramedullary nailing in patients of advanced age. Poole et al.[15] and Whitmarsh et al.[16] used three-dimensional CT imaging to develop cortical thickness maps of the proximal femur for evaluation of osteoporosis treatments. Turmezei et al.[17] used three-dimensional CT imaging to describe bone changes observed in hip osteoarthritis. Such methods have yet to be applied to MRI of the hip in vivo.

The goal of this study was to develop a three-dimensional MRI-based method to assess cortical bone thickness, an important contributor to bone strength and fracture risk. An MRI-based tool would have the advantages of higher resolution compared and no ionizing radiation dose compared to CT. We describe the development of the tool and assess its ability to reproducibly detect regional changes in the proximal femur. Finally, we applied the tool in vivo to 89 subjects' proximal femurs. We hypothesized that it would be possible to detect regional differences in cortical thickness in the proximal femur in these subjects.

Materials and Methods

Human Subjects

This study had institutional review board approval and written informed consent was obtained from all subjects.

MR Imaging

All MR imaging was performed on a 3T MRI scanner (Siemens Skyra, Erlangen, Germany) using a 26-element coil setup composed of flexible 18-element array coil anteriorly and 8 elements from a spine coil posteriorly (Siemens, Erlangen, Germany). Each hip was imaged using a three-dimensional fast low angle shot sequence (TR = 37 ms, TE = 4.92 ms, matrix 512×512, field of view = 12 cm, slice thickness 1.5 mm, 60 coronal images) with a generalized auto calibrating partially parallel acquisition (GRAPPA) at acceleration factor of two (scan time = 15 min and 18 s).

Segmentation of MR Images

The workflow for generation of the cortical thickness maps is provided in Figure 1. First, the freely available FireVoxel software package (NYU Center for Advanced Imaging Innovation and Research, New York, USA) was used to perform the manual segmentations for all cases. Each image set was manually segmented to include two regions of interest: 1. the trabecular compartment of the proximal femur and 2. the whole proximal femur (Fig. 1).

Surface Generation and Parcellation

The 3DSlicer v4.4 toolkit (National Alliance of Medical Image Computing, NA-MIC) was used to generate and smooth three-dimensional surface representations of the trabecular compartment and the whole proximal femur using the marching cubes algorithm, 10 iterations of sinc smoothing, and 25% surface point decimation.

The ParaView v4.3.1 (Kitware, New York, USA) visualization software was used to parcellate the proximal femur (both the whole femur and trabecular compartment) into five regions: 1. femoral head, 2. femoral neck, 3. greater trochanter, 4. intertrochanteric region, and 5. subtrochanteric region (Fig. 2).

Three-Dimensional Cortical Mapping

Finally, custom software was written using the C++ programming language and visualization toolkit (VTK, Kitware, New York, USA) to calculate the Euclidean distance[18] between the trabecular compartment and the outer cortex of the proximal femur, representing the cortical bone thickness. Mean cortical thickness values for each region was calculated. In addition, a color map representing the cortical thickness was created for each proximal femur.

Rater Reliability Evaluation

Sample Subset and Rater Qualifications—To demonstrate the reliability of MRI manual bone segmentation, a subset of 10 randomly selected MR image sets were selected for further analysis. Each subject was manually segmented twice by four different raters with varying experience levels: 1. expert, 2. physician, 3. medical student, and 4. undergraduate student. Segmentation volume, image overlap metrics, and resulting finite element model stiffness were calculated to compare the segmentations between raters.

Image Overlap Metrics—We also evaluated image segmentation performance by computing standard image overlap metrics.[19] These were calculated using the C++ programming language and the Insight Toolkit (ITK). These metrics include the union overlap (Jaccard coefficient), total overlap (sensitivity), mean overlap (Dice coefficient), volume similarity, false negative, and false positive (1-specificity). Perfect matching segmentations are represented by a value of 1 for the union overlap, total overlap, and mean overlap, and a value of 0 for false negatives, false positive, and volume similarity.

Statistical Analysis—IBM SPSS Statistics software (IBM, Armonk, New York, USA) was used to perform statistical analyses. A t-test was used to compare differences in mean cortical thickness between the femoral head, femoral neck, greater trochanter, intertrochanteric region, and subtrochanteric region between the osteoporotic proximal femur subjects. A p-value < 0.05 was considered statistically significant.

To assess reproducibility and reliability, we computed the intrarater (i.e., for each of the four raters) and an aggregate interrater (i.e., over all four raters) median coefficients of variation (CV) and the intraclass correlation coefficients (ICC), as described by Shrout and Fleiss.[20] As per Landis and Koch,[21] ICC reliability values can be classified as follows: slight (0 to 0.2), fair (0.21 to 0.4), moderate (0.41 to 0.6), substantial (0.61 to 0.80), and almost perfect (0.81 to 1.0).

Results

Eighty-nine subjects were recruited for this study with an average age of 61 years. Three-dimensional regional cortical thickness maps were generated for all 89 subjects. A typical

three-dimensional segmentation from our dataset is demonstrated in Figure 3 and a sample three-dimensional cortical thickness map is demonstrated in Figure 4. The average cortical thickness in different regions of the proximal femur ranged from 1.1 mm (femoral head) to 3.1 mm (subtrochanteric region). Table 1 summarizes the average cortical thickness measurements for each region. There was a statistically significant difference between regions for all possible comparisons ($p < 0.01$). Table 2 demonstrates the p-values for each regional comparison.

With regards to the reliability of our segmentations, Table 3 provides the whole proximal femur volumes for the 10 subjects for both segmentations of the four users. Table 4 provides the intrarater and aggregate interrater root-meansquare coefficients of variation for whole femur volume. We report the median coefficient of variation of 0.53% over all four raters with respect to segmented volume. Table 5 provides the intrarater and aggregate interrater intraclass correlation coefficients for whole femur volume. The intrarater and aggregate interrater intraclass correlation coefficients for whole femur volume were all 0.99.

Table 6 demonstrates the intrarater study based on the image overlap metrics. The intrarater mean values of 0.970, 0.947, and 0.973 correspond to the total overlap, union overlap, and mean overlap metrics, respectively. The intrarater mean values of −0.005, 0.030, and 0.025 correspond to the volume similarity, false negative, and false positive metrics, respectively. Table 7 demonstrates the interrater study based on the image overlap metrics. The interrater mean values of 0.967, 0.935, and 0.966 correspond to the total overlap, union overlap, and mean overlap metrics, respectively. The interrater mean values of 0.001, 0.033, and 0.034 correspond to the volume similarity, false negative, and false positive metrics, respectively.

Discussion

The dual energy x-ray absorptiometry (DXA) scan is the standard-of-care method for diagnosing and monitoring osteoporosis; however, DXA incompletely captures the three-dimensional properties of bone structure that contribute to fracture risk. Others have demonstrated the use of cortical thickness maps using CT imaging for investigating the effect of osteoporosis drug therapy.[15,16] Magnetic resonance imaging does not expose the patient to ionizing radiation required by CT. Furthermore, MR imaging allows for higher resolution of microstructures that are not possible with CT due to the latter’s ionizing radiation dose limitations and photon detector width in modern scanners.

In this study, we have demonstrated the ability to generate three-dimensional cortical bone thickness maps for the proximal femur from non-contrast high resolution MRI, which advantageously does not require ionizing radiation and offers higher spatial resolution compared to CT. This image processing tool can be implemented on any desktop computer, and we further used it to parcellate the proximal femur to into different regions; the tool was able to detect regional differences in cortical thickness within the proximal femur. The latter is important because regional variations in bone quality, fracture risk, and therapy response are likely to exist within the proximal femur given that, for example, the femoral neck and greater trochanter are the sites most likely to fracture. In the future, the combination of trabecular microstructure analysis and cortical thickness mapping may offer more accurate

methods to predict fragility fractures in patients or to monitor disease progression and treatment response longitudinally. In the future, we aim to study larger cohorts of patients to investigate whether cortical mapping could be used in combination with trabecular microstructure analysis or DXA to more accurately predict osteoporotic fractures and monitor therapy response.

Furthermore, we demonstrated that manual segmentation of MR images of the proximal femur can be performed with high reproducibility and reliability within raters and between raters of different training levels. This is important for large cohort studies and longitudinal studies. Using different methods to comprehensively evaluate the segmentation performance, we have demonstrated that the segmentations were reproducibly and reliably performed by the same rater and between the raters (all trained in the image anatomy by a musculoskeletal radiologist) despite being in different stages of their medical education. Overall, the high reproducibility validates the use of multiple manual raters for clinical applications where large numbers of image datasets and segmentations need to be generated by different users in order to maintain clinical throughput.

By comparing the segmentation volumes using the coefficient of variation, we see that all raters ranging from the expert rater to the medical student and undergraduate raters demonstrated excellent segmentation reproducibility with median within-rater CVs less than 1%. Furthermore, the between-rater median CV was also less than 1%. Overall, this result means that there is essentially no difference in the variation between segmentations performed within a rater or between raters. This is reinforced by the high calculated ICC values (> 0.9) which can be classified using Shrout and Fleiss as “almost perfect” reliability when tested within and between raters.[20]

Using the six commonly reported image overlap metrics, we have demonstrated that the segmentations were a close match not only within the raters but between them as well. When compared to manual segmentations performed using CT imaging of the phalanx bones, proximal tibia, and distal femur, our overlap metric values are better than reported for the phalanx bones[2] and similar to those reported for the knee joint.[22] In comparison to other MRI studies that have compared manual raters of the proximal femur, our results are similar but demonstrate slightly better overlap characteristics.[23]

This study focused on manual segmentation as a method of identifying bone from MR images, which is currently the gold standard for bone segmentation from MR imaging. Ideally, the process of identifying the proximal femur will be fully automatic (or at least semi-automatic), which as previously mentioned has been attempted. To reach the accuracy of manual segmentation, most automated segmentation techniques still require manual editing, especially in regions of thin cortical bone, close cortical contact, and pathologic anatomy. Future progress toward fully automated segmentation, potentially taking advantage of machine learning methods, will foster expedited image analysis for large clinical studies.

This study has limitations. First, the images of proximal femur microarchitecture are lower in resolution compared to images of bone microarchitecture previously obtained in the distal tibia or distal radius, either with MRI[9] or high-resolution peripheral computed tomography

scanners (HR-pQCT).[24] However, the hip is a clinically important site affected in osteoporosis and osteoarthritis, and only recently has hip microarchitecture imaging become possible via MRI due to improvements in radiofrequency coil technology that boosts the signal-to-noise ratio. We note that the images are higher in resolution than those obtained by standard clinical hip CT or hip MR imaging and that HRpQCT scanners cannot image the axial skeleton. Second, the scan time is relatively long at 15 minutes. Subtle patient movement could also introduce variation in measurements that cannot be corrected by extremely accurate and precise manual segmentation. Retrospective motion correction methods may help alleviate this problem in the future.[9] Finally, we note that this study was performed only at one site, and in the future it will be important that multi-center studies are performed to demonstrate that the method can be fully translated to the clinical setting. In the future, it may be possible to use web-based training to teach anatomy and the segmentation software (FireVoxel or others) to raters at different institutions or even implement machine learning methods for semi-automatic and automatic segmentation.

Conclusions

This study describes the development of a novel image processing tool to generate MRI-based cortical thickness maps of the proximal femur. Advantageously, this tool can be implemented on any desktop computer. Furthermore, compared to CT, MRI offers the advantages of no ionizing radiation administration and higher resolution. Finally, we have demonstrated the robustness of our techniques with high measurement reproducibility and reliability both within and between raters of different levels of experience. Given the importance of cortical bone thickness for bone strength and fracture risk, this tool would permit the safe and reliable serial monitoring of the effectiveness of different drug therapies on proximal femur cortical thickness in osteoporosis clinical trials. Additionally, these three-dimensional cortical thickness maps could be used in combination with other bone quality assessment methods, such as finite element analysis, microarchitectural analysis, or even BMD measurements, to determine the best predictors of osteoporotic fracture.

Acknowledgments

This project was supported in part by grants from the National Institutes of Health; NIH R01-AR070131, NIH R01-AR066008, and NIH R01 AR0683282.

References

1. Johannesdottir F, Turmezei T, Poole KE. Cortical bone assessed with clinical computed tomography at the proximal femur. J Bone Miner Res. 2014;29(4):771–83. [PubMed: 24677244]
2. DeVries NA, Gassman EE, Kallemeyn NA, et al. Validation of phalanx bone three-dimensional surface segmentation from computed tomography images using laser scanning. Skeletal Radiol. 2008;37(1):35–42. [PubMed: 17962937]
3. Chang G, Honig S, Brown R, et al. Finite Element Analysis Applied to 3-T MR imaging of proximal femur microarchitecture: lower bone strength in patients with fragility fractures compared with control subjects. Radiology. 2014;272(2):464–74. [PubMed: 24689884]
4. Rong Q, Bai J, Huang Y, Lin J. Biomechanical assessment of a patient-specific knee implant design using finite element method. Biomed Res Int. 2014;2014:353690. [PubMed: 25101275]

5. Crowley AR, Dong J, McHaffie A, et al. Measuring bone erosion and edema in rheumatoid arthritis: a comparison of manual segmentation and RAMRIS methods. J Magn Reson Imaging. 2011;33(2):364–71. [PubMed: 21274978]
6. Emond PD, Inglis D, Choi A, et al. Volume measurement of bone erosions in magnetic resonance images of patients with rheumatoid arthritis. Magn Reson Med. 2012;67(3):814–23. [PubMed: 21688319]
7. Wehrli FW, Hwang SN, Ma J, et al. Cancellous bone volume and structure in the forearm: noninvasive assessment with MR microimaging and image processing. Radiology. 1998;206(2):347–57. [PubMed: 9457185]
8. Majumdar S, Link TM, Augat P, et al. Trabecular bone architecture in the distal radius using magnetic resonance imaging in subjects with fractures of the proximal femur. Magnetic Resonance Science Center and Osteoporosis and Arthritis Research Group. Osteoporos Int. 1999;10(3):231–9. [PubMed: 10525716]
9. Wehrli FW. Structural and functional assessment of trabecular and cortical bone by micro magnetic resonance imaging. J Magn Reson Imaging. 2007;25(2):390–409. [PubMed: 17260403]
10. Krug R, Banerjee S, Han ET, et al. Feasibility of in vivo structural analysis of high-resolution magnetic resonance images of the proximal femur. Osteoporos Int. 2005;16(11):1307–14. [PubMed: 15999292]
11. Chang G, Deniz CM, Honig S, et al. Feasibility of three-dimensional MRI of proximal femur microarchitecture at 3 tesla using 26 receive elements without and with parallel imaging. J Magn Reson Imaging. 2014;40(1):229–38. [PubMed: 24711013]
12. Burge R, Dawson-Hughes B, Solomon DH, et al. Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. J Bone Miner Res. 2007;22(3):465–75. [PubMed: 17144789]
13. Yelin E, Murphy L, Cisternas MG, et al. Medical care expenditures and earnings losses among persons with arthritis and other rheumatic conditions in 2003, and comparisons with 1997. Arthritis Rheum. 2007;56(5):1397–407. [PubMed: 17469096]
14. Tucker D, Surup T, Petersik A, Kelly M. Full circle: 3D femoral mapping demonstrates age-related changes that influence femoral implant positioning. Injury. 2016 2;47(2):471–7. [PubMed: 26686594]
15. Poole KE, Treece GM, Gee AH, et al. Denosumab rapidly increases cortical bone in key locations of the femur: a 3D bone mapping study in women with osteoporosis. J Bone Miner Res. 2015;30(1):46–54. [PubMed: 25088963]
16. Whitmarsh T, Treece GM, Gee AH, Poole KE. Mapping Bone Changes at the Proximal Femoral Cortex of Postmenopausal Women in Response to Alendronate and Teriparatide Alone, Combined or Sequentially. J Bone Miner Res. 2015;30(7):1309–18. [PubMed: 25639838]
17. Turmezei TD, Treece GM, Gee AH, et al. Quantitative 3D analysis of bone in hip osteoarthritis using clinical computed tomography. Eur Radiol. 2016 7;26(7):2047–54. [PubMed: 26443603]
18. Danielsson PE. Euclidean distance mapping. Computer Graphics and Image Processing. 1980;14(3):227–48.
19. Tustison N Introducing Dice, Jaccard, and Other Label Overlap Measures to ITK. Insight Journal 2009 Jul-Dec, Available at: http://insight-journal.org/browse/publication/707.
20. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8. [PubMed: 18839484]
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics, 1977;33(1):159–74. [PubMed: 843571]
22. Ramme AJ, Criswell AJ, Wolf BR, et al. EM segmentation of the distal femur and proximal tibia: a high-throughput approach to anatomic surface generation. Ann Biomed Eng. 2011;39(5):1555–62. [PubMed: 21222162]
23. Zoroofi RA, Sato Y, Nishii T, Sugano N, et al. Automated segmentation of necrotic femoral head from 3D MR data. Comput Med Iimasging Graph. 2004;28(5):267–78.
24. Patsch JM, Burghardt AJ, Kazakia G, Majumdar S. Noninvasive imaging of bone microarchitecture. Ann N Y Acad Sci. 2011;1240:77–87. [PubMed: 22172043]

Figure 1.
Manual segmentation of MRI of the proximal femur: **A**, trabecular compartment, **B**, outer cortex, and **C**, combined trabecular and outer cortex.

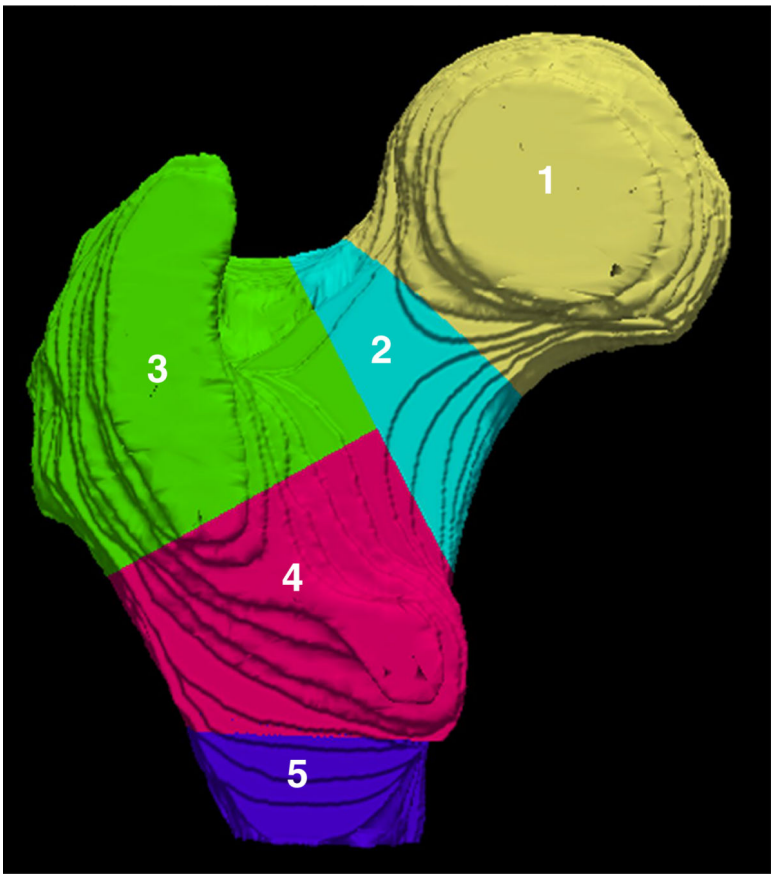

Figure 2.
Surface parcellation of the proximal femur into five regions: femoral head (1, yellow), femoral neck (2, aqua), greater trochanter (3, green), intertrochanteric region (4, pink), and subtrochanteric region (5, purple).

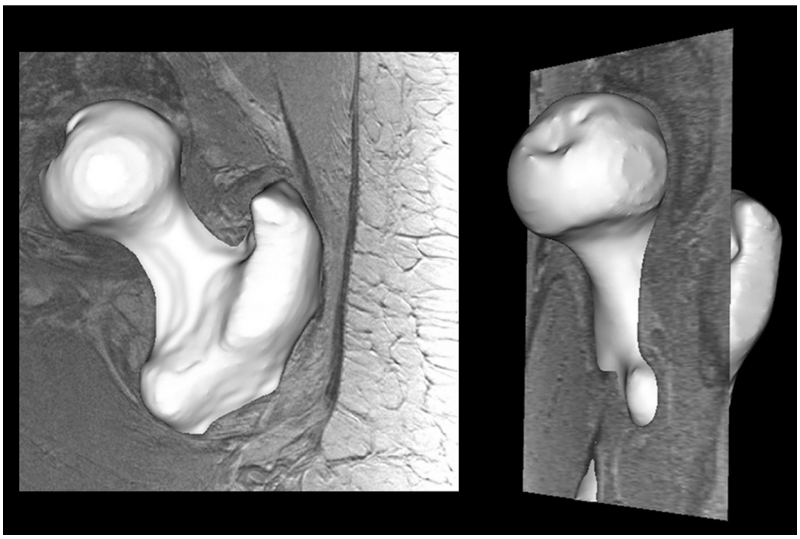

Figure 3.
Typical three dimensional segmentation result from our dataset: coronal view (left), sagittal view (right).

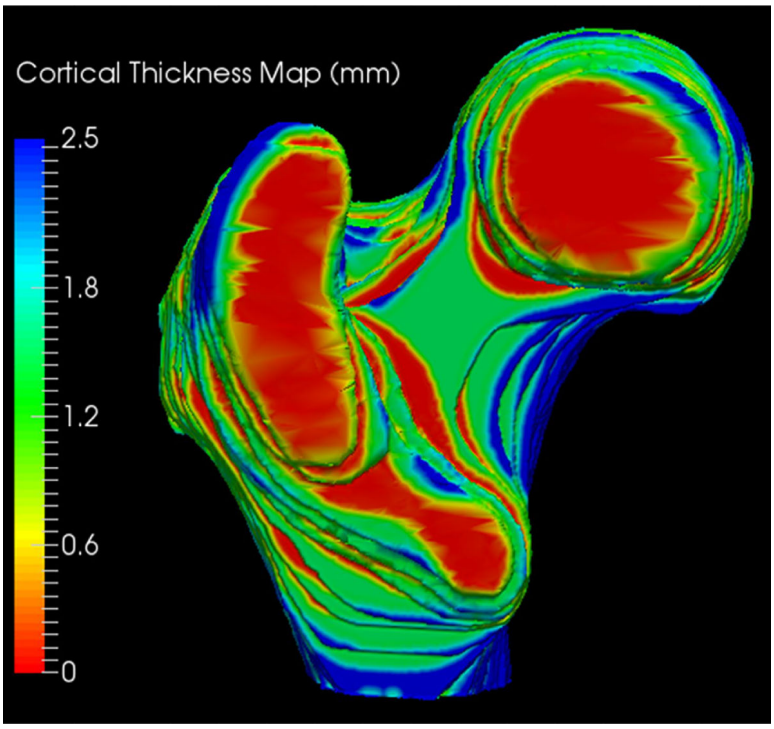

Figure 4.
Sample three dimensional cortical thickness map of the proximal femur.

Table 1

Average Cortical Thickness and Standard Deviations Based on Parcellated Region of the Proximal Femurs

Region of Interest	Average Cortical Thickness (mm)
Whole Proximal Femur	1.697 (0.295)
Femoral Head	1.123 (0.217)
Femoral Neck	1.944 (0.385)
Intertrochanteric Region	2.114 (0.328)
Greater Trochanter	1.202 (0.206)
Subtrochanteric Region	3.126 (0.638)

Table 2

P-values from Statistical t-testing Comparing the Regions of the Proximal Femurs of the 89 Subjects

	Femoral Head	Femoral Neck	Intertrochanteric Region	Greater Trochanter	Subtrochanteric Region
Femoral Head		5.39×10^{-40}	3.37×10^{-56}	0.014	5.63×10^{-65}
Femoral Neck			1.92×10^{-3}	7.79×10^{-36}	3.02×10^{-32}
Intertrochanteric Region				2.42×10^{-52}	1.06×10^{-27}
Greater Trochanter					6.91×10^{-63}
Subtrochanteric Region					

Bull Hosp Jt Dis (2013). Author manuscript; available in PMC 2020 July 06.

Table 3

Calculated Femur Volumes (cm^3) from Both Segmentations for Each Rater for the 10 Cases

	Expert Rater		Physician Rater		Medical Student Rater		Undergraduate Rater	
Case 1	104.71	106.33	102.74	102.54	106.89	106.43	102.39	106.01
Case 2	128.39	128.24	126.38	127.33	129.73	126.95	127.22	129.87
Case 3	196.81	195.72	194.67	193.94	197.72	195.55	192.12	198.18
Case 4	137.97	137.77	134.97	136.25	142.37	139.39	135.29	136.95
Case 5	115.05	115.51	113.18	114.30	116.57	117.34	116.99	117.20
Case 6	108.03	106.49	103.24	102.66	103.83	104.57	103.46	104.17
Case 7	123.54	121.01	118.76	118.42	114.81	122.78	121.85	122.51
Case 8	169.70	171.62	167.11	168.35	171.13	171.89	167.66	168.40
Case 9	116.77	117.68	113.45	114.57	115.99	115.59	117.05	119.22
Case 10	126.60	125.66	123.42	127.04	127.87	127.08	125.17	126.65

Table 4

Median Coefficients of Variation (CV) and Interquartile Ranges for Each of the Four Raters and Over All Four Raters (Aggregate) for Assessment of Whole Femur Volume

	Expert Rater	Physician Rater	Medical Student Rater	Undergraduate Rater	Aggregate
Whole Femur Volume	0.53% (0.23%–1.0%)	0.52% (0.24%–0.69%)	0.48% (0.31%–1.5%)	0.85% (0.36%–1.6%)	0.53% (0.31%–1.1%)

Bull Hosp Jt Dis (2013). Author manuscript; available in PMC 2020 July 06.

Table 5

Intraclass Correlation Coefficients (ICC) with Confidence Intervals for Each of the Four Raters and Over All Four Raters (Aggregate) for Assessment of Whole Femur Volume

	Expert Rater	Physician Rater	Medical Student Rater	Undergraduate Rater	Aggregate
Whole Femur Volume	0.99 (0.99–1.0)	0.99 (0.99–1.0)	0.99 (0.97–1.0)	0.99 (0.99–1.0)	0.99 (0.99–1.0)

Bull Hosp Jt Dis (2013). Author manuscript; available in PMC 2020 July 06.

Table 6

Overall Intrarater Overlap Metrics for Four Different Raters Each Performing Each Segmentation Twice

Metric	Minimum Value	Maximum Value	Average Value	Standard Deviation
Total Overlap	0.911	0.989	0.970	0.013
Union Overlap	0.890	0.969	0.947	0.014
Mean Overlap	0.942	0.984	0.973	0.008
Volume Similarity	−0.067	0.022	−0.005	0.017
False Negative	0.011	0.089	0.030	0.013
False Positive	0.011	0.049	0.025	0.008

Table 7

Overall Interrater Overlap Metrics for Four Different Raters Each Performing Two Segmentations Per Case

Metric	Minimum Value	Maximum Value	Average Value	Standard Deviation
Total Overlap	0.902	0.993	0.967	0.015
Union Overlap	0.872	0.964	0.935	0.018
Mean Overlap	0.932	0.982	0.966	0.010
Volume Similarity	−0.065	0.073	0.001	0.023
False Negative	0.007	0.098	0.033	0.015
False Positive	0.006	0.096	0.034	0.014

Bull Hosp Jt Dis (2013). Author manuscript; available in PMC 2020 July 06.