Research Article

Segmentation of Hip Cartilage in Compositional Magnetic Resonance Imaging: A Fast, Accurate, Reproducible, and Clinically Viable Semi-Automated Methodology[†]

Author List

1. Dr Scott Fernquest[1] : BA BMBCh, scott.fernquest@ndorms.ox.ac.uk, Botnar Research Centre, Oxford, Old Road, OX3 7LD, Tel: +44 (0)1865 227313. Study design, acquisition and analysis of data, manuscript drafting and review. Corresponding Author.

2. Dr Daniel Park[1]. Study design, analysis of data, manuscript drafting and review.

3. Dr Marija Marcan[2]. Acquisition and analysis of data, manuscript drafting and review.

4. Mr Antony Palmer[1]. Study design, manuscript drafting and review.

5. Dr Irina Voiculescu[2]. Study design, acquisition and analysis of data, manuscript drafting and review.

6. Prof Sion Glyn-Jones[1]. Study design, manuscript drafting and review.

All authors have read and approved the final submitted manuscript.

Ethical approval for this study was granted by Oxfordshire Research Ethics Committee B (07Q1605/26).

[1]Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal sciences, University of Oxford, Oxford, United Kingdom, OX3 7LD

[2]Department of Computer Science, University of Oxford, Oxford, United Kingdom, OX1 3QD

Received 19 October 2017; Revised 21 December 2017; Accepted 16 February 2018
Journal of Orthopaedic Research

DOI 10.1002/jor.23881

Abstract

Manual segmentation is a significant obstacle in the analysis of compositional MRI for clinical decision-making and research. Our aim was to produce a fast, accurate, reproducible, and clinically viable semi-automated method for segmentation of hip MRI. We produced a semi-automated segmentation method for cartilage segmentation of hip MRI sequences consisting of a two step process: (1) fully automated hierarchical partitioning of the data volume generated using a bespoke segmentation approach applied recursively, followed by (2) user selection of the regions of interest using a region editor. This was applied to dGEMRIC scans at 3T taken from a prospective longitudinal study of individuals considered at high risk of developing osteoarthritis (SibKids) which were also manually segmented for comparison. Fourteen hips were segmented both manually and using our semi-automated method. Per hip, processing time for semi-automated and manual segmentation was 10-15 minutes, and 60-120 minutes respectively. Accuracy and Dice similarity coefficient (DSC) for the comparison of semi-automated and manual segmentations was 0.9886 and 0.8803 respectively. Intra-observer and inter-observer reproducibility of the semi-automated segmentation method gave an accuracy of 0.9997 and 0.9991, and DSC of 0.9726 and 0.9354 respectively. We have proposed a fast, accurate, reproducible, and clinically viable semi-automated method for segmentation of hip MRI sequences. This enables accurate anatomical and biochemical measurements to be obtained quickly and reproducibly. This is the first such method that shows clinical applicability, and could have large ramifications for the use of compositional MRI in research and clinically.

Keywords

1. Compositional MRI
2. Hip
3. Segmentation

Introduction

This article is protected by copyright. All rights reserved

At present, the only effective treatment for osteoarthritis (OA) of large weight bearing joints, such as the hip and knee, is joint replacement surgery for end stage disease. The prolonged period of conservative management for pain and loss of function prior to surgery results in a large socioeconomic burden, costing between 1% and 2.5% of the United Kingdom's gross domestic product [1].

An increasing number of disease modifying pharmaceutical[2] and surgical early intervention[3] strategies for OA have been proposed. However, these require the ability to identify OA early and detect small changes in cartilage structure and function to enable treatment efficacy to be evaluated within an acceptable timeframe. This is not possible using conventional imaging techniques such as plain film radiography which only allow OA diagnosis once irreversible loss of cartilage has occurred. Radiographic assessment of OA relies mainly on the evaluation of both osteophytes and joint space narrowing [4; 5], which are late changes and lack soft tissue depiction.

Musculoskeletal Magnetic Resonance Imaging (MRI) is a rapidly advancing imaging methodology[6]. MRI allows qualitative and quantitative assessment of soft tissue and joint anatomy. MRI can detect cartilage loss when none is evident on plain radiography [7; 8], thus offering early diagnostic potential for OA [9]. More recently Compositional MRI techniques have been proposed that offer the potential to diagnose pre-structural osteoarthritis by evaluating the biochemical properties of tissues[9]. Delayed Gadolinium Enhanced Magnetic Resonance Imaging of Cartilage (dGEMRIC) is the best-validated technique and indirectly measures cartilage glycosaminoglycan content[10]. This permits more accurate quantitative assessment, allowing the recognition of the earliest signs of

OA within cartilage, before even any change in morphological measurements are seen [9]. Accordingly, compositional MRI is a current focus of OA research [11].

To use MRI data in this capacity, however, requires the articular cartilage first be isolated from neighbouring structures, i.e. segmented [12]. Segmentation allows the analysis of specific anatomical components of the joint[13], and the generation of 3D models[14]. Both of which are valuable tools in the pre-operative planning and treatment stratification of patients undergoing joint replacement or joint preservation surgery for conditions such as Femoroacetabular Impingement (FAI)[15]. Segmentation, however, represents a large obstacle for the translation of promising compositional MRI techniques into clinical practice [16] and large cohort studies [6].

The accuracy of segmentation significantly influences the error and precision of the analysis of the desired joint component. However, due to the structure and morphology of the articular cartilages, as well as the nature of MR acquisition, obtaining accurate segmentations can be problematic [17]. MR scans can be manually segmented slice by slice by experts. Yet for routine clinical use and in large clinical trials manual methods are too time consuming, requiring months of experience, training in both segmentation methodology and manual segmentation software, and taking a number of hours per joint [17] which significantly elevates cost. Moreover, they are prone to inter- and intra-observer variability and subjectivity of results. Thus before compositional MRI can be widely used both in large trials and clinically there is a need to automate the segmentation method. The main challenges in developing an automatic method are the

thin structure of the cartilage and the low contrast both between the femoral and acetabular cartilages and with the surrounding soft tissues [18].

Recently, several successful semi-automatic and automatic methods have been proposed for cartilage segmentation in the knee, including machine based learning approaches [18; 19], atlas based approaches [20; 21], model based approaches such as Active Shape Modelling (ASM) [17], and texture analysis [22]. These have allowed quantitative analysis of knee cartilage biochemistry and morphology both in a slice-by-slice manner and as a 3D representation. In contrast, there have been very few successful examples of semi-automated or automated segmentation in the hip. The Partial Volume Effect (PVE) seen at the close interface of the femoral and acetabular cartilage, caused by the spherical nature of the joint, and thinner cartilage than in the knee [23], has resulted in difficulty in achieving a successful and clinically applicable automated hip segmentation tool. A number of authors have proposed methodologies, however issues such as the requirement of traction, the need for a training data set, and long processing times have severely limited their usefulness [24; 25 26 13 27 28; 29].

We aimed to produce a semi-automated method of segmentation of MRI of the hip that is fast, accurate, reproducible, and clinically viable, that could allow wider implementation of compositional MRI in clinical and research arenas.

Methodology

Ethical approval for this study was granted by Oxfordshire Research Ethics Committee B (07Q1605/26).

Population

Participants were selected from a prospective longitudinal study of individuals considered at high risk of developing osteoarthritis (SibKids) [30 31 32]. The dGEMRIC scans used in this study were taken from the 5-year follow up point [33]. Sibkids and spouse controls were selected for dGEMRIC of a single hip if both hips fulfilled the criteria: (1) No investigation or treatment for hip pain within the previous two years; (2) Minimum joint space width >2.5mm and Kellgren-Lawrence Grade <2 on AnteroPosterior (AP) pelvis radiographs; and (3) No radiographic evidence of dysplasia or pincer morphology on AP or lateral pelvis radiographs. Compositional MRI is likely to be of most use in individuals with conditions predisposing to OA, such as cam-type FAI, but who do not show significant radiographic evidence of disease[9]. Accordingly, individuals were selected from the SibKid cohort for this study if they displayed Cam morphology but had no significant radiographic disease progression, defined as a reduction in minimum joint space width greater than 0.5mm[34].

Radiograph Protocol

A standing Anteroposterior and cross-table lateral radiograph was acquired. Anteroposterior radiographs were performed with the beam centred in the midline between the superior border of the pubic symphysis and a line connecting the anterior superior iliac spines[35]. Radiographs were repeated if the obturator foramen index was outside 0.7-1.4[36]. Lateral radiographs were performed with the beam parallel to the table and orientated 45 degrees to the index hip centred on the femoral head. The hip was positioned in 15 degrees of internal rotation with the aid of a wedge[35]. Radiographs

were analysed using OxMorf 2.1.0-dev6 software in a non-sequential manner by two observers. Cam and pincer morphology was evaluated on the anteroposterior and cross-table lateral radiograph using the alpha angle[37] and centre-edge angle. Cam morphology was defined as an alpha angle greater than 60 degrees on anteroposterior radiographs[38]. Pincer morphology was defined as centre-edge angle greater than 39 degrees on AP and lateral radiographs[39].

MRI Protocol

MRI scans were performed using a 3 Tesla Philips Achieva X-series (Philips Healthcare, Eindhoven, Netherlands) platform with two flexible surface coils (medium and large). 0.2mM/Kg of Magnevist (dimeglumine gadopentetate [Gd-DTPA2-], Bayer Schering Pharma, Germany) was administered intravenously. An exercise protocol was completed with 10 minutes of walking on a treadmill at 4km/hour followed by 150 hip movements (50 flexion, 50 internal rotation, 50 external rotation) to ensure full perfusion of the gadolinium into the articular cartilage[40]. 75 minutes after contrast administration the dGEMRIC sequence was commenced. Sequence parameters comprised sagittal inversion-prepared 3D-turbo-field-echo (TFE) with repetition time (TRTFE) 6.0ms, echo time (TE) 2.9ms, flip angle 12 degrees, bandwidth 289Hz/pixel, inversion times (Tis) 2100, 1200, 600, 250, and 105ms, field of view 180mm x 180mm, slice thickness 3mm, acquisition matrix 208 x 209 (interpolated to 512 x 512). The first slice was aligned with the most medial aspect of the femoral head and the remaining slices extending laterally with no gap between slices. To attain sufficient signal-to noise at short Tis, the total time between inversion pulses (TRTOTAL) was held constant at 2200ms. Scan time was

45 minutes. Quantitative T1 maps were generated by averaging signal intensity from segmented areas on co-registered images and fitting a mono-exponential T1 recovery curve using a non-linear algorithm (MATLAB, MA, USA).

Manual Segmentation

Sagittal dGEMRIC images were manually segmented using OsiriX Software (Version 6.0.2 64 Bit, Pixmeo, Geneva, Switzerland). A single observer (academic musculoskeletal clinician) manually segmented the 1200 ms inversion time images to isolate the articular cartilage of the joint. Regions of interest (ROI) were developed based on a clockface around the centre of the femoral head at 30° intervals as previously described [13]. Femoral and acetabular cartilage was segmented and regions of interest were not extended beyond where there was adjacent femoral and acetabular cartilage, hence didn't include the chondrolabral junction.

Semi-automated segmentation

The semi-automated segmentation method consists of a two-step process: (1) fully automated hierarchical partitioning of the data volume generated using a bespoke segmentation approach applied recursively [41], followed by (2) user selection of the anatomical structure of interest using a bespoke region editor.

The fully automated partitioning regards the image as a terrain map, in which the intensity of each voxel represents its altitude in the terrain. Within this metaphor, local minima are voxels surrounded by voxel of higher intensity. Each image is processed simulating a flood that starts from such local minima and fills the terrain further uphill. The places where two catchment water basins meet are retained as

relevant contours. This process is applied recursively, until the whole image is flooded into a single 'lake'. Snapshots of meeting contours generated in this way are taken at different times during the process, thus creating a hierarchy of layers of contours which partition the image. The lowest layer of the hierarchy consists of the voxels in the source image. At each stage of the recursion a coarser layer of regions is generated such that each region contains sub-regions present in the layer below (Figure 1).

Since different kinds of tissue have different scan intensity signatures, the gradients between areas of interest will be emphasised by this process. The boundaries between regions mould themselves around relevant boundaries between tissues. Although this process is easiest described for a single image slice, the algorithm is oblivious to the dimensionality of the voxel collection. In practice, if the inter-slice thickness is small, the segmentation is carried out directly on a full 3D volume. If the inter-slice thickness is larger than 3mm the process is more accurate if the segmentation is carried out in a slice-by-slice fashion. Regardless of the dimensionality of the segmentation algorithm, the hierarchical partitioning output is interacted with via slice-by-slice 2D views in one of the scan planes. Prior to the partitioning through flooding, a one-off calibration step is applied in order to determine an appropriate windowing level (i.e. window centre and window width of the range of relevant intensities) for the specific MRI sequence. The calibration is carried out manually by adjusting sliding controls on the user interface until the relevant boundaries feature a good contrast. The calibration takes at most 5 minutes, and its settings are then retained for any other scans of the same type. The calibration step only needs to be performed once irrespective of the number of scans it is applied to, so long as the

scanning and sequence parameters are kept constant.

Once the partition hierarchy is in place, user selection of the anatomical structure of interest proceeds using a bespoke region editor. The user selects anatomical regions that align with the tissue of interest, such as the articular cartilage. This is carried out by selecting the appropriate layer in which most of the tissue is selected appropriately, and clicking the relevant region(s). Each selected region can subsequently be refined across different coarseness layers in the hierarchical partitioning such that sub regions can be selected or unselected (Figure 2). The selection of an anatomical region of interest, in this case the cartilage, can span across several layers of the hierarchy [42; 43].

Further interaction can take place through contour drawing tools to refine the segmentation if needed. A bespoke data structure and associated algorithms have been designed for the purpose of manipulating and storing image hierarchies generated in this way. The region selection process is specific to viewing a slice at a time, although if the segmentation has been carried out in 3D then clicking a region in a slice selects the voxel volume corresponding to that region across all the slices in which it is present.

At the end of this process, results from a collection of slices can be embedded into a single 3D region of interest. Segmentation thus produces a mask that provides volumetric and relaxation time data for the selected region. This mask can be overlaid on the image viewing or exported in conventional data formats, so that its voxels can be compared directly against gold standard selections (such as manual

segmentations), counting the number of true positives (T_P), true negatives (T_N), false positives (F_P), and false negatives (F_N).

Statistical Analysis

The semi-automated cartilage segmentations were compared to gold standard binary manual segmentations on a pixel-by-pixel basis using a custom script in MATLAB R2014b (Mathworks, Natick, MA, USA) using several conventional metrics such as: sensitivity = $T_P/(T_P + F_N)$, specificity = $T_N/(T_N + F_P)$, accuracy = $(T_P + F_N)/(T_P + T_N + F_P = F_N)$, and Dice similarity coefficient (DSC) = $2TP/(2T_P + F_P + F_N)$. The sensitivity is the "true positive fraction" and specificity the "true negative fraction," while DSC is a spatial overlap index [17; 44]. Intra- and inter-observer reproducibility of manual and semi-automated segmentation was measured similarly, but with the gold standard being the original segmentation and the comparator being the repeat segmentation. A much wider selection of metrics exists for comparing segmentation results[45]. This is because each are sensitive to different aspects of the comparison, hence none of them can be declared as definitive comparison metric. We have chosen the above set as relevant for comparing our results with similar results in the literature.

Results

Semi-automated versus manual segmentation

Twenty-five individuals had interpretable dGEMRIC scans at follow-up. Of these, fourteen individuals had cam morphology and did not show any evidence of disease progression on radiographs. The articular cartilage from these fourteen hips was

segmented using the dGEMRIC 1200 ms inversion time image both manually and using our semi-automated methodology. Each scan gave approximately 15 sagittal slices through the hip. The number of slices per hip from which articular cartilage was segmented varied between patients depending on hip size and image quality but was consistently between 4 and 6. A single academic musculoskeletal clinician (SF) trained in segmentation performed both the manual and semi-automated segmentations. Processing time for manual segmentation ranged on average between 60 and 90 minutes per hip. The automated pre-processing component of our proposed method, which produces the hierarchical partitioning, ranged between 5 and 10 minutes per hip. The user selection component, i.e. the selection of the anatomical structure of interest, took a newly trained operator on average 5 minutes per hip. This gave an overall time range for our semi-automatic segmentation method of between 10 and 15 minutes. The operator training time amounted to a one-off 15-minute period. In the comparison of the masks produced from the semi-automated segmentations with the manual segmentations the accuracy was 0.9886, with a sensitivity of 0.9418 and specificity of 0.9984, and a DSC of 0.8803 (Table 1).

Reproducibility

Five hips were chosen at random from the cohort to undergo additional manual segmentation, and all fourteen underwent semi-automated segmentation, of articular cartilage by the first academic musculoskeletal clinician (SF) more than 4 weeks after the first segmentation and by a second academic musculoskeletal clinician trained in segmentation of the hip (AP). This allowed us to determine and compare the intra-observer and inter-observer reproducibility of both the manual and semi-automated

methodologies. Only five hips underwent repeat manual segmentation as the reproducibility of this methodology has been extensively validated in the literature[33; 46] and was not the primary aim of this paper. The time intensive nature of data acquisition from manual segmentation also played a role in selecting only five hips for repeat manual analysis. For manual segmentation intra-observer reproducibility showed an accuracy of 0.9992 and DSC of 0.9410, and inter-observer reproducibility showed an accuracy of 0.9987 and DSC 0.9036 respectively. For semi-automated segmentation intra-observer reproducibility showed an accuracy of 0.9997 and DSC of 0.9726, and inter-observer reproducibility showed an accuracy of 0.9991 and DSC of 0.9354 (Table1).

Discussion

Magnetic Resonance imaging (MRI) offers the potential of diagnosing OA at a stage where patients may benefit from intervention, and acting as an assay of disease to test the efficacy of novel early intervention treatments [9]. However, manual segmentation is time consuming and a viable automated segmentation method in the hip remains elusive. Accordingly, segmentation presents itself as a large challenge in the translation of promising compositional MRI techniques into clinical practice and large cohort studies [6 16]. We have proposed a fast, accurate, reproducible, and clinically viable semi-automated segmentation method that may allow wider implementation of compositional MRI for use in quantitative analysis of early OA in the hip.

Volumetric (3D) hip-dGEMRIC is inhibited in its biochemical quantitative analysis of

the hip joint by the complexity involved in visualising and segmenting the articular cartilage, typically restricting evaluation to a few slices. Domayer et al. [13] manually reformatted an isotropic 3D T1 dataset to evaluate cartilage quality at multiple Regions Of Interest (ROIs) by using the clockface method, allowing clinical interpretation. A method we employed in our manual segmentations. Regardless, the time intensive nature of the manual segmentation in this process makes it unsuitable for use in larger trials or clinically.

The average computation time for the automated part of the method in our study was 5-10 minutes on a computer with a single 3.40 GHz Core i5-3570 (Intel, Santa Clara, California, USA). The average time necessary for user selection post computation was 5 minutes for each hip. Thus giving an overall processing time for our semi-automated segmentation method of 10-15 minutes per hip. This is significantly faster than our own manual segmentation time of 60-90 minutes, and that quoted by Pollard et al. [33] of 70-100 minutes, using the same manual segmentation method, for each hip. Moreover, with the reduced number of user decisions and input the level of training required to use our semi-automated segmentation method is significantly less than that required for manual segmentation.

Our semi-automated segmentation time is comparable to the processing times for semi-automated and automated segmentation in the knee given by Folkesson et al. [18] of 10 minutes. Our processing time is similar to that proposed by Chandra et al. [29] of 10 minutes and considerably quicker than that stated for automated segmentation of the hip by Siversson et al. [27] of 3 hours in their work on automated segmentation of the hip. It is worth noting however that both authors required the use of a training data

set of manual segmentations to initialize the automated segmentation method, limiting their widespread applicability. Moreover, Siversson et al. required the use of 20 Intel Xeon cores to achieve their fully automated approach. It is also worth noting that other authors who have published automated and semi-automated methodology for segmentation of the hip such as Nishii et al. [25], Li et al. [26], Sato et al. [47], and Xia et al. [28] have made no mention of the processing time or processing power required for their methodology. Whilst each of these segmentations was carried out on a different dataset and using different computing power, it nevertheless makes sense to compare the order of magnitude of the time taken and the (extensive or otherwise) prerequisites required by each method.

Our semi-automated segmentation methodology showed high levels of accuracy when validated against manual segmentations of the same hip, with an overall accuracy of 0.9886 (±0.0315) and a DSC of 0.8803 (±0.0211). This level of accuracy is better than that reported by Siverrson et al. [27] in their fully automated segmentation of the hip methodology, where DSC for cartilage was 0.824 (±0.052). It is also superior to that reported by Xia et al. [28] where DSC was 0.81 (±0.03) for combined cartilage volumes in the hip when compared against manual segmentation. Our level of accuracy is comparable to that proposed by Fripp et al. [17] where DSC was 0.89 for combined cartilage volumes in the knee when compared against manual segmentation.

The method of semi-automated segmentation proposed in this paper showed high inter-observer reproducibility with a DSC 0.9354. Most importantly, this was higher than the inter-observer reproducibility seen in the manual segmentation of the same hips, where DSC was 0.9036.

Overall, the results strongly support the use of our proposed methodology as an accurate, fast, and reproducible method of semi-automated segmentation in the hip. The potential ramifications of this work are two fold, affecting both clinical and research settings. This rapid method of hip segmentation would (a) greatly increase the viability, and (b) reduce the cost, of large scale trials using compositional MRI, that are so often hampered by the man hours required to obtain quantitative data from each hip scan. Clinically, compositional MRI is performed in the pre-operative assessment of patients for hip arthroscopy and hip preservation surgery, with an increasing number of arthroscopic hip procedures performed each year worldwide [48]. It is not widely used, however, due to its time consuming nature and on-going validation. Yet, the speed, accuracy, and reproducibility with which our method allows such compositional MRI analysis of patients greatly increases the clinical applicability and utilization of compositional MRI.

A limitation of this automated segmentation methodology is that it does not distinguish between acetabular and femoral cartilage throughout the joint, and does not split the cartilage into individual regions of interest (ROIs). This is not surprising considering that historically the greatest difficulty in any automated hip segmentation methodology has come from trying to determine the acetabular and femoral cartilage interface. Moreover, Domayer et al. (10) reported not being able to make that distinction regularly in their manual methodology. The creation of ROIs is not, however, unachievable with our current methodology. Our bespoke segmentation method is capable of determining the acetabular and femoral cartilage interface in the better-delineated images, and the

creation of ROIs based on a clockface could form part of the manual process or be coded into the automated process.

Whether automated or manual we do not envisage that production of ROIs would significantly affect the accuracy of the morphological results, as it would simply be a further division of the already segmented cartilage within the predefined boundaries of the mask. Nor do we believe it would significantly increase the processing time of the semi-automated method. The addition of these features is something we intend to establish as part of our methodology in further work to optimize our semi-automated segmentation technique.[41-43] This would allow the biochemical and morphological analysis of specific areas of cartilage between patients and within patients over time.

Further limitations of this work include the lack of biochemical comparison data obtained between the manual and semi-automated segmentations. Given the accuracy of the morphological comparison, and lack of subdivided ROIs currently, we felt that comparison of biochemical data would prove of no further benefit at this point. Moreover, although the training period and overall time taken to segment each hip is reported as being very short, it should be noted that these timings are for academic musculoskeletal clinicians previously trained in manual segmentation. These timings may be longer in an untrained individual, though introducing this methodology to untrained users will form a component of the next stage of this research.

It should also be noted that the methodology in this study has only been validated on individuals with cam-type FAI. Primarily this was because it is not currently thought

that pincer-type FAI predisposes to OA[49], and so compositional MRI would be of little use in this condition. Dysplasia is a well-recognised risk factor for the development of OA, and compositional MRI has been shown to play a role in the diagnosis of early OA changes in dysplastic individuals[50]. However, this disease process was outside of the scope of the current study, as such this methodology would require further validation before being used in dysplastic individuals. Moreover, this study has only compared segmentations obtained semi-automatically to the gold standard of those obtained manually from a single scan performed using the sequence parameters as described above. The effect of patient positioning, joint unloading, or changes in sequence parameters on the accuracy of this methodology were not investigated.

Conclusion

In conclusion this study has proposed a fast, accurate, and reproducible semi-automated method for segmentation of MRI in the hip. This methodology is capable of providing both morphological and biochemical quantitative data from MRI sequences. Our methodology allows hips to be segmented an order of magnitude faster than previously possible manually. Accordingly, this methodology could have large ramifications for the scale at which compositional MRI is used in both research and clinically. However, what is clear from this study is the need for further optimization and research to create a method capable of producing individual ROIs and defining the interface between the acetabular and femoral cartilage, which would allow quantitative analysis of specific cartilage subregions. This would have a significant impact on the use of compositional MRI sequences in a research setting.

Regardless, this study forwards the argument for the widespread use of compositional MRI as an accurate and efficient clinical tool to select appropriate candidates for hip arthroscopy and hip preservation surgery, and as a research tool capable of large-scale use

Acknowledgements

We would like to acknowledge funding support from the National Institute for Health Research (NIHR) Oxford Musculoskeletal Biomedical Research Unit**.**

We would like to thank Claudio Pereira and Prof Neal Bangerter for their invaluable assistance with participant assessments and MRI sequences.

Disclosures

Authors Scott Fernquest, Daniel Park, Marija Marcan, Antony Palmer, and Irina Voiculescu have nothing to disclose.

Sion Glyn-Jones receives payment for lectures including service on speakers bureaus from Zimmer-Biomet, Corin, and ConMed.

Figure Legends

Figure 1. Representation of hierarchical partitioning of dGEMRIC scan of hip in sagittal view. A) Layer 0. B) Layer 1. C) Layer 2. D) Layer 3. E) Layer 4. F) Layer 5

Figure 2 Selection of anatomical structure of interest using semi-automated segmentation methodology. A and B) Selection of cartilage from hierarchical partitioning across different layers. C) Area of cartilage selection from semi-automated segmentation overlaying original image MRI image.

Table Legends

Table 1. Average and Standard Deviation values for Dice Similarity Coefficient, Accuracy, Sensitivity, and Specificity for manual versus semi-automated segmentation of cartilage anatomy; comparison and reproducibility.

References

- 1. Hiligsmann M, Cooper C, Arden N, et al. 2013. Health economics in the field of osteoarthritis: an expert's consensus paper from the European Society for Clinical and Economic Aspects of Osteoporosis and Osteoarthritis (ESCEO). Semin Arthritis Rheum 43:303-313.
- 2. Hunter DJ. 2011. Pharmacologic therapy for osteoarthritis--the era of disease modification. Nat Rev Rheumatol 7:13-22.
- 3. Palmer AJ, Malak TT, Broomfield J, et al. 2016. Past and projected temporal trends in arthroscopic hip surgery in England between 2002 and 2013. BMJ Open Sport Exerc Med 2:e000082.

- 4. Altman R, Asch E, Bloch D, et al. 1986. Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association. Arthritis and rheumatism 29:1039-1049.
- 5. Kellgren JH, Lawrence JS. 1957. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 16:494-502.
- 6. Roemer FW, Crema MD, Trattnig S, et al. 2011. Advances in imaging of osteoarthritis and cartilage. Radiology 260:332-354.
- 7. Amin S, LaValley MP, Guermazi A, et al. 2005. The relationship between cartilage loss on magnetic resonance imaging and radiographic progression in men and women with knee osteoarthritis. Arthritis and rheumatism 52:3152-3159.
- 8. Guermazi A, Niu J, Hayashi D, et al. 2012. Prevalence of abnormalities in knees detected by MRI in adults without knee osteoarthritis: population based observational study (Framingham Osteoarthritis Study). BMJ 345:e5339.
- 9. Palmer AJ, Brown CP, McNally EG, et al. 2013. Non-invasive imaging of cartilage in early osteoarthritis. The bone & joint journal 95-B:738-746.
- 10. Bashir A, Gray ML, Burstein D. 1996. Gd-DTPA2- as a measure of cartilage degradation. Magnetic resonance in medicine 36:665-673.
- 11. Conaghan PG, Hunter DJ, Maillefert JF, et al. 2011. Summary and recommendations of the OARSI FDA osteoarthritis Assessment of Structural Change Working Group. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 19:606-610.

- 12. Eckstein F, Cicuttini F, Raynauld JP, et al. 2006. Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): morphological assessment. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 14 Suppl A:A46-75.
- 13. Domayer SE, Mamisch TC, Kress I, et al. 2010. Radial dGEMRIC in developmental dysplasia of the hip and in femoroacetabular impingement: preliminary results. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 18:1421-1428.
- 14. De Momi E, Pavan E, Motyl C, et al. 2005. Hip joint anatomy virtual and stereolithographic reconstruction for preoperative planning of total hip replacement. CARS 2005: Computer Assisted Radiology and Surgery Proceedings of the 19th International Congress and Exhibition 1281:708-712.
- 15. Leunig M, Beaule PE, Ganz R. 2009. The concept of femoroacetabular impingement: current status and future perspectives. Clinical orthopaedics and related research 467:616-622.
- 16. Pedoia V, Majumdar S, Link TM. 2016. Segmentation of joint and musculoskeletal tissue in the study of arthritis. Magma 29:207-221.
- 17. Fripp J, Crozier S, Warfield SK, et al. 2010. Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. IEEE transactions on medical imaging 29:55-64.
- 18. Folkesson J, Dam EB, Olsen OF, et al. 2007. Segmenting articular cartilage automatically using a voxel classification approach. IEEE transactions on medical imaging 26:106-115.
- 19. Prasoon A, Igel C, Loog M, et al. 2013. Femoral cartilage segmentation in knee MRI scans using two stage voxel classification. Conference

proceedings : Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual Conference 2013:5469-5472.

- 20. Carballido-Gamio J, Majumdar S. 2011. Atlas-based knee cartilage assessment. Magnetic resonance in medicine 66:574-583.
- 21. Shan L, Zach C, Charles C, et al. 2014. Automatic atlas-based three-label cartilage segmentation from MR knee images. Medical image analysis 18:1233-1246.
- 22. Dodin P, Pelletier JP, Martel-Pelletier J, et al. 2010. Automatic human knee cartilage segmentation from 3D magnetic resonance images. IEEE transactions on bio-medical engineering 57.
- 23. Cheng Y, Guo C, Wang Y, et al. 2013. Accuracy limits for the thickness measurement of the hip joint cartilage in 3-D MR images: simulation and validation. IEEE transactions on bio-medical engineering 60:517-533.
- 24. Nakanishi K, Tanaka H, Sugano N, et al. 2001. MR-based three-dimensional presentation of cartilage thickness in the femoral head. European radiology 11:2178-2183.
- 25. Nishii T, Sugano N, Sato Y, et al. 2004. Three-dimensional distribution of acetabular cartilage thickness in patients with hip dysplasia: a fully automated computational analysis of MR imaging. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 12:650-657.
- 26. Li W, Abram F, Beaudoin G, et al. 2008. Human hip joint cartilage: MRI quantitative thickness and volume measurements discriminating acetabulum and femoral head. IEEE transactions on bio-medical engineering 55:2731-2740.

- 27. Siversson C, Akhondi-Asl A, Bixby S, et al. 2014. Three-dimensional hip cartilage quality assessment of morphology and dGEMRIC by planar maps and automated segmentation. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 22:1511-1515.
- 28. Xia Y, Chandra SS, Engstrom C, et al. 2014. Automatic hip cartilage segmentation from 3D MR images using arc-weighted graph searching. Physics in medicine and biology 59:7245-7266.
- 29. Chandra SS, Surowiec R, Ho C, et al. 2016. Automated analysis of hip joint cartilage combining MR T2 and three-dimensional fast-spin-echo images. Magnetic resonance in medicine 75:403-413.
- 30. Chitnavis J, Sinsheimer JS, Clipsham K, et al. 1997. Genetic influences in end-stage osteoarthritis. Sibling risks of hip and knee replacement for idiopathic osteoarthritis. The Journal of bone and joint surgery British volume 79:660-664.
- 31. Spencer JM, Loughlin J, Clipsham K, et al. 2005. Genetic background increases the risk of hip osteoarthritis. Clinical orthopaedics and related research:134-137.
- 32. Pollard TC, Batra RN, Judge A, et al. 2012. Genetic predisposition to the presence and 5-year clinical progression of hip osteoarthritis. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 20:368-375.
- 33. Pollard TC, McNally EG, Wilson DC, et al. 2010. Localized cartilage assessment with three-dimensional dGEMRIC in asymptomatic hips with normal morphology and cam deformity. The Journal of bone and joint surgery American volume 92:2557-2569.

- 34. Dawson J, Fitzpatrick R, Carr A, et al. 1996. Questionnaire on the perceptions of patients about total hip replacement. The Journal of bone and joint surgery British volume 78:185-190.
- 35. Clohisy JC, Carlisle JC, Beaule PE, et al. 2008. A systematic approach to the plain radiographic evaluation of the young adult hip. The Journal of bone and joint surgery American volume 90 Suppl 4:47-66.
- 36. Jacobsen S, Sonne-Holm S, Lund B, et al. 2004. Pelvic orientation and assessment of hip dysplasia in adults. Acta Orthop Scand 75:721-729.
- 37. Notzli HP, Wyss TF, Stoecklin CH, et al. 2002. The contour of the femoral head-neck junction as a predictor for the risk of anterior impingement. The Journal of bone and joint surgery British volume 84:556-560.
- 38. Agricola R, Waarsing JH, Thomas GE, et al. 2014. Cam impingement: defining the presence of a cam deformity by the alpha angle: data from the CHECK cohort and Chingford cohort. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 22:218-225.
- 39. Tannast M, Siebenrock KA, Anderson SE. 2007. Femoroacetabular impingement: radiographic diagnosis--what the radiologist should know. AJR Am J Roentgenol 188:1540-1552.
- 40. Burstein D, Velyvis J, Scott KT, et al. 2001. Protocol issues for delayed Gd(DTPA)(2-)-enhanced MRI (dGEMRIC) for clinical evaluation of articular cartilage. Magnetic resonance in medicine 45:36-41.
- 41. Marcan M, Voiculescu I. 2016. Unsupervised segmentation of MRI knees using image partition forests. SPIE Medical Imaging SPIE International Society for Optics and Photonics.

- 42. Golodetz S, Nicholls D, Voiculescu I, et al. 2014. Two tree-based methods for the waterfall. Pattern Recognition 47:3276-3292.
- 43. Golodetz S, Voiculescu I, Cameron S. 2017. Simpler editing of graph-based segmentation hierarchies using zipping algorithms. Pattern Recognition 70:44-59.
- 44. Dice LR. 1945. Measures of the Amount of Ecologic Association between Species. Ecology 26:297-302.
- 45. Yeghiazaryan V, Voiculescu I. 2017. Boundary overlap for medical image segmentation evaluation. Proc Spie 10135.
- 46. Palmer A, Fernquest S, Rombach I, et al. 2017. Diagnostic and prognostic value of delayed Gadolinium Enhanced Magnetic Resonance Imaging of Cartilage (dGEMRIC) in early osteoarthritis of the hip. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society.
- 47. Sato R, Takeyama H, Tanaka T, et al. 2001. Development of high-performance and rapid immunoassay for model food allergen lysozyme using antibody-conjugated bacterial magnetic particles and fully automated system. Applied biochemistry and biotechnology 91-93:109-116.
- 48. AJR Palmer TM, L Broomfield, J Holton, L Majkowski, GER Thomas, A Taylor, AJ Andrade, G Collins, K Watson, AJ Carr, S Glyn-Jones. 2016. Past and projected temporal trends in arthroscopic hip surgery in England between 2002 and 2013. BMJ Open Sport & Exercise medicine 2:82-87.
- 49. Saberi Hosnijeh F, Zuiderwijk ME, Versteeg M, et al. 2017. Cam Deformity and Acetabular Dysplasia as Risk Factors for Hip Osteoarthritis. Arthritis Rheumatol 69:86-93.

- 50. Kim YJ, Jaramillo D, Millis MB, et al. 2003. Assessment of early osteoarthritis in hip dysplasia with delayed gadolinium-enhanced magnetic resonance imaging of cartilage. The Journal of bone and joint surgery American volume 85-A:1987-1992.

Tables

Table 1: Average and Standard Deviation values for Dice Similarity Coefficient, Accuracy, Sensitivity, and Specificity for manual versus semi-automated segmentation of cartilage anatomy; comparison and reproducibility.

Segmentation		Dice Similarity Coefficient (DSC)		Accuracy		Sensitivity		Specificity	
		Average	SD	Average	SD	Average	SD	Average	SD
Manual vs. Semi-automated		0.8803	0.0211	0.9886	0.0315	0.9418	0.0232	0.9984	0.0015
Semi-automated vs. Semi-automated	Intra-observer	0.9726	0.0093	0.9997	0.0009	0.9808	0.0183	0.9996	0.0003
	Inter-Observer	0.9354	0.0231	0.9991	0.0004	0.9009	0.0551	0.9998	0.0003
Manual vs. Manual	Intra-observer	0.9410	0.0142	0.9992	0.0001	0.9796	0.0115	0.9993	0.0001
	Inter-Observer	0.9036	0.0141	0.9987	0.0002	0.9660	0.0204	0.9990	0.0002

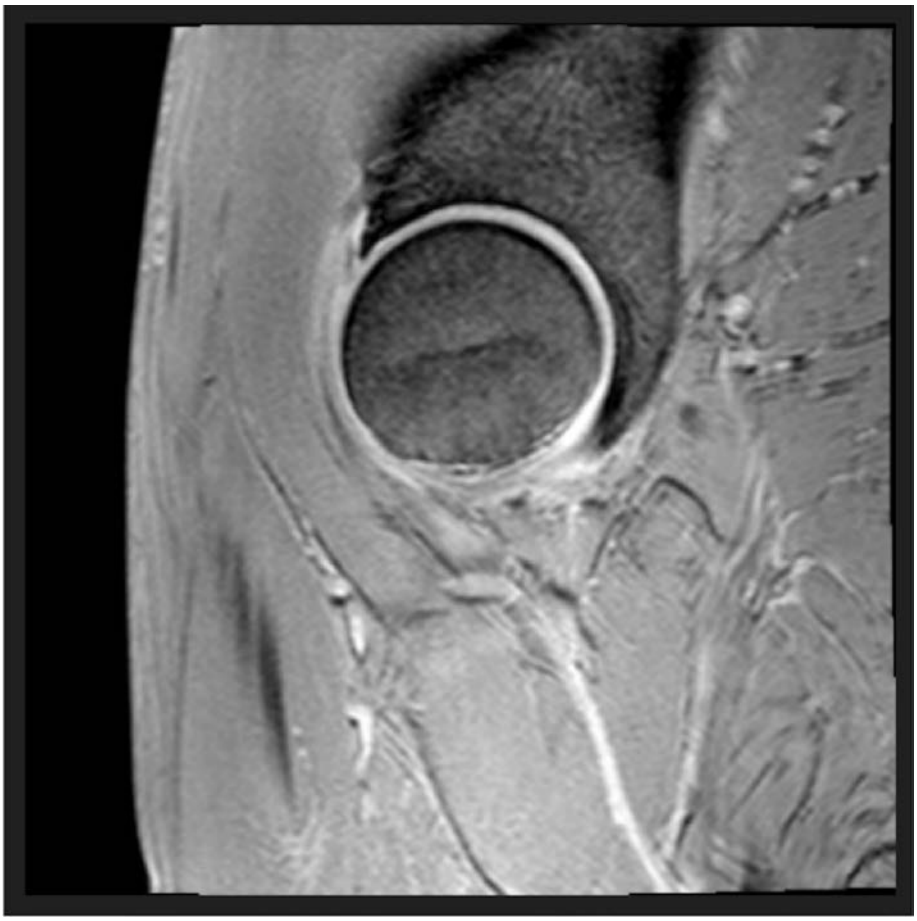

Figure 1a

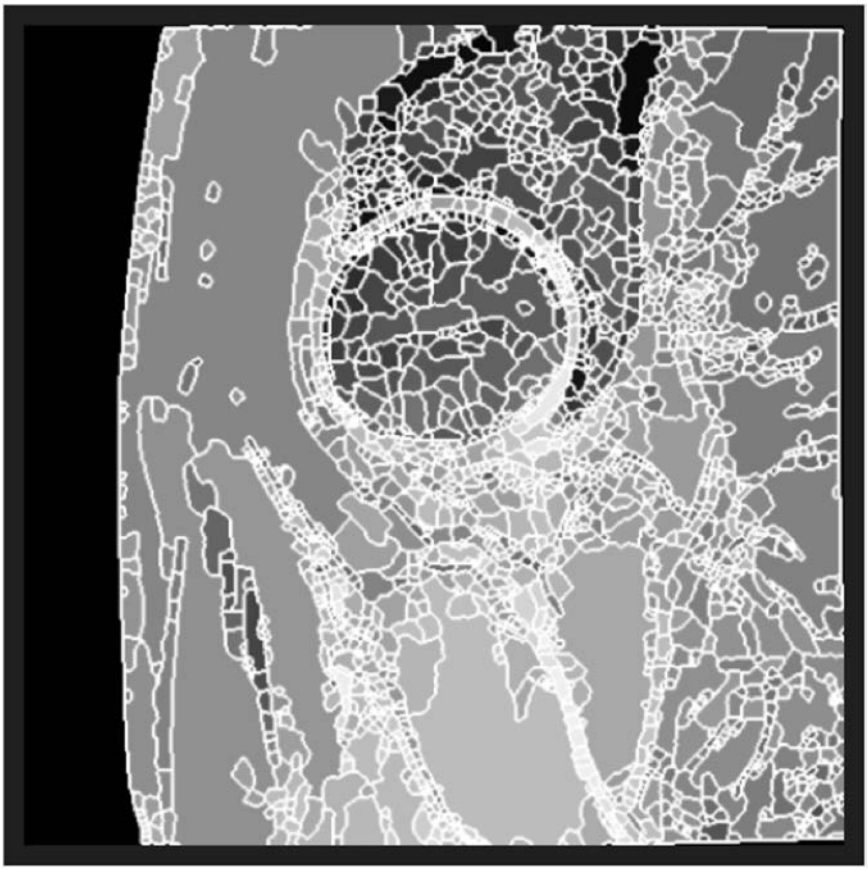

Figure 1b

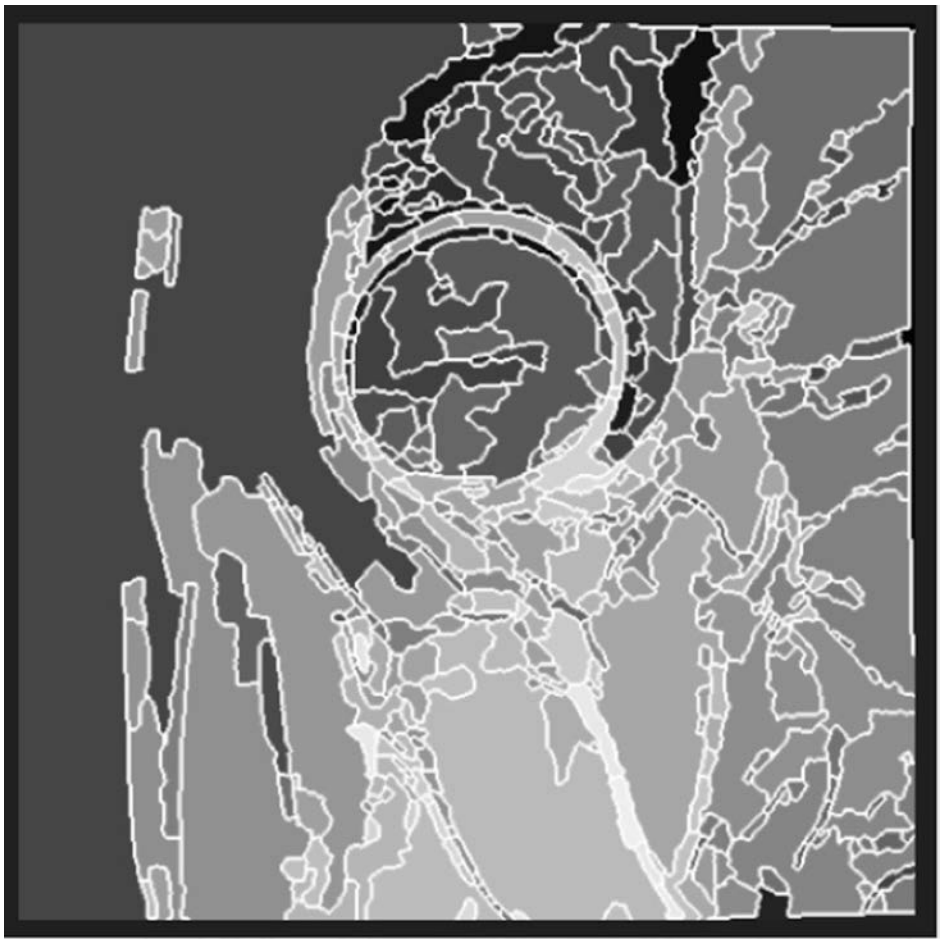

Figure 1c

Figure 1d

Figure 1e

Figure 1f

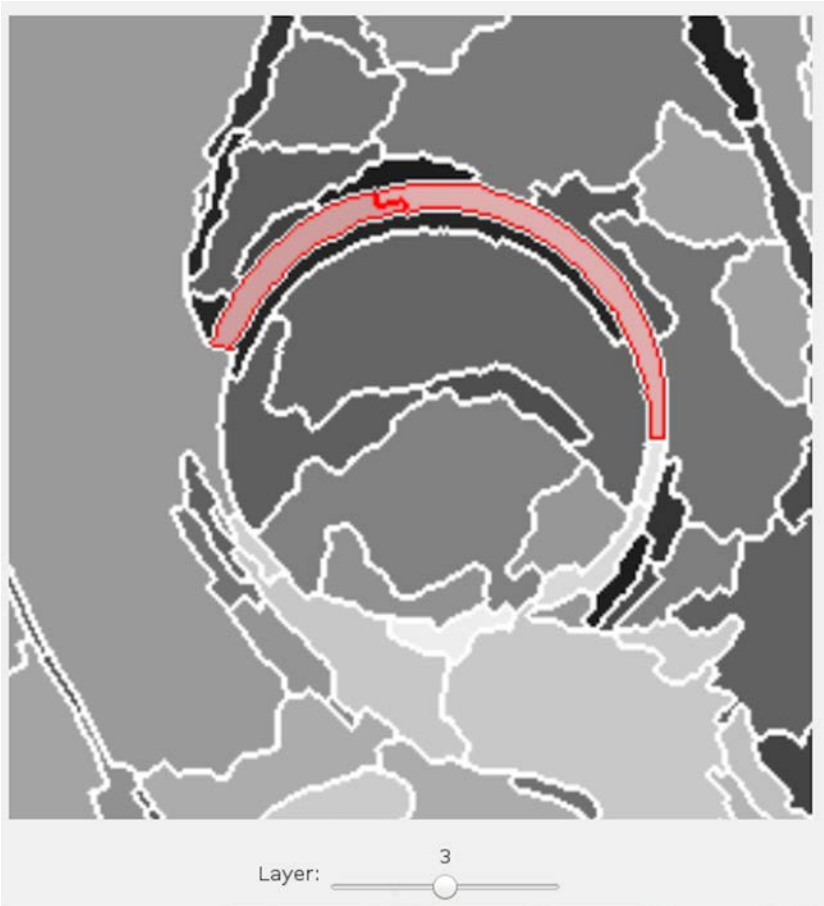

Figure 2a

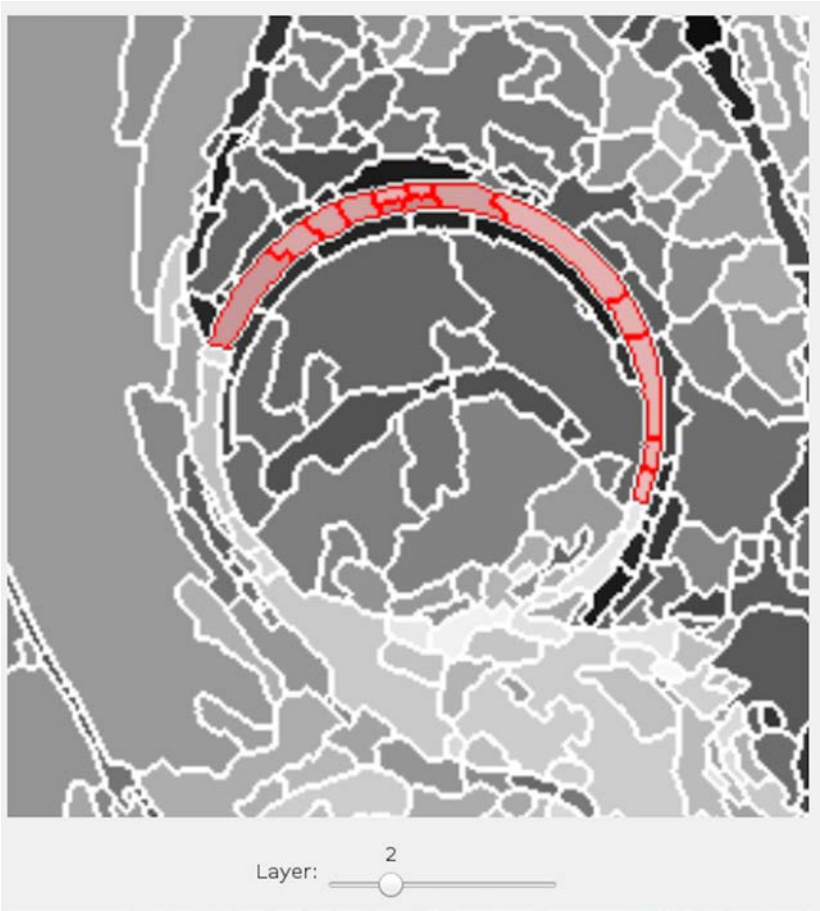

Figure 2b

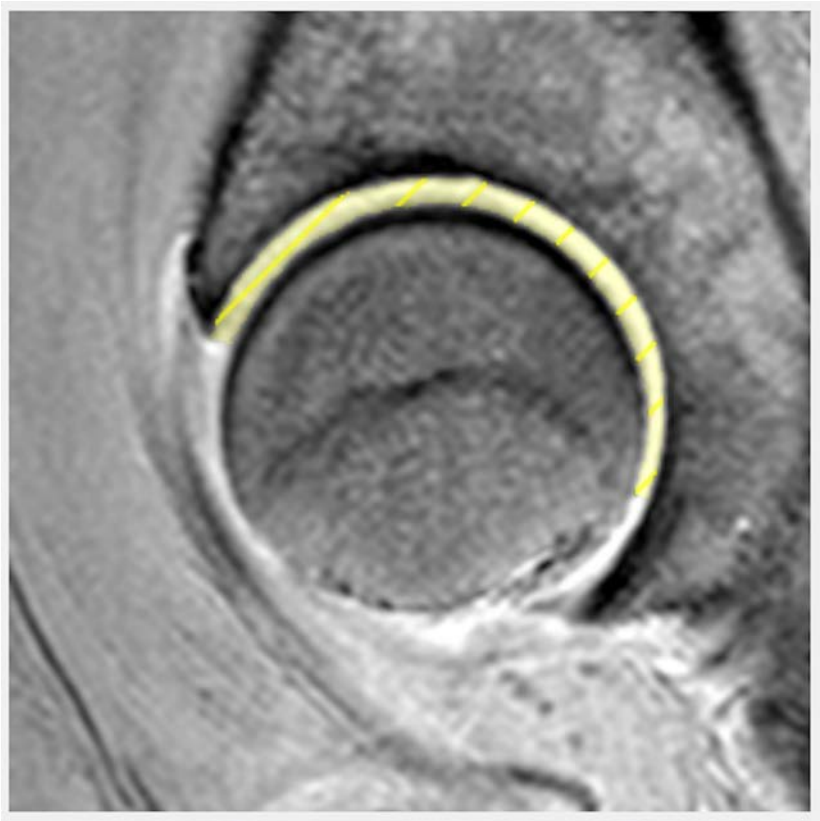

Figure 2c