



# Validation of watershed-based segmentation of the cartilage surface from sequential CT arthrography scans

Mary E. Hall<sup>1</sup>, Marianne S. Black<sup>1,2</sup>, Garry E. Gold<sup>2,3</sup>, Marc E. Levenston<sup>1,2,3</sup>

<sup>1</sup>Department of Mechanical Engineering, Stanford University, Stanford, CA, USA; <sup>2</sup>Department of Radiology, Stanford University, Stanford, CA, USA; <sup>3</sup>Department of Bioengineering, Stanford University, Stanford, CA, USA

**Contributions:** (I) Conception and design: All authors; (II) Administrative support: None; (III) Provision of study materials or patients: ME Levenston, GE Gold; (IV) Collection and assembly of data: ME Hall; (V) Data analysis and interpretation: ME Hall, ME Levenston; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

**Correspondence to:** Marc E. Levenston. 452 Escondido Mall, Room 225, Stanford, CA 94305, USA. Email: [levenston@stanford.edu](mailto:levenston@stanford.edu).

**Background:** This study investigated the utility of a 2-dimensional watershed algorithm for identifying the cartilage surface in computed tomography (CT) arthrograms of the knee up to 33 minutes after an intra-articular iohexol injection as boundary blurring increased.

**Methods:** A 2D watershed algorithm was applied to CT arthrograms of 3 bovine stifle joints taken 3, 8, 18, and 33 minutes after iohexol injection and used to segment tibial cartilage. Thickness measurements were compared to a reference standard thickness measurement and the 3-minute time point scan.

**Results:** 77.2% of cartilage thickness measurements were within 0.2 mm (1 voxel) of the thickness calculated in the reference scan at the 3-minute time point. 42% fewer voxels could be segmented from the 33-minute scan than the 3-minute scan due to diffusion of the contrast agent out of the joint space and into the cartilage, leading to blurring of the cartilage boundary. The traced watershed lines were closer to the location of the cartilage surface in areas where tissues were in direct contact with each other (cartilage-cartilage or cartilage-meniscus contact).

**Conclusions:** The use of watershed dam lines to guide cartilage segmentation shows promise for identifying cartilage boundaries from CT arthrograms in areas where soft tissues are in direct contact with each other.

**Keywords:** Computed tomography (CT); arthrography; knee; image processing; segmentation

Submitted Sep 15, 2020. Accepted for publication Jul 12, 2021.

doi: [10.21037/qims-20-1062](https://doi.org/10.21037/qims-20-1062)

**View this article at:** <https://dx.doi.org/10.21037/qims-20-1062>

## Introduction

Osteoarthritis is a debilitating and expensive joint disease that affects a large portion of the adult population worldwide (1). Currently, joint space narrowing as shown by a standing radiograph is the clinical gold standard for detecting osteoarthritis, but this does not allow for early detection or assessment of cartilage health. Imaging biomarkers may provide a way to detect joint diseases such as knee osteoarthritis at an early stage in a minimally invasive way.

Computed tomography (CT) and magnetic resonance imaging (MRI) have been used in various studies to detect morphology (2,3), structure (4,5), and content of molecules such as collagen and glycosaminoglycans (GAGs) in articular cartilage (6-8). Though MRI enables non-invasive assessment of cartilage, some patients are not good candidates for MRI exams due to circumstances such as medical device implants, claustrophobia, or inability to stay motionless for long periods of time. In these cases, contrast enhanced CT, where an iodinated contrast agent is

injected intra-articularly into the joint, is an alternative to MRI for assessment of cartilage. In addition, because CT scans are generally faster and less expensive than MRI, they can be valuable for capturing instantaneous morphology of cartilage at multiple time points.

This ability is particularly useful for techniques such as weight bearing CT imaging of the knee (9,10), which aim to detect small morphological changes in soft tissues such as articular cartilage compression over time. New developments in motion correction (11-17) have enabled imaging of weight bearing subjects with higher image quality. Weight bearing CT has been used previously to detect joint space width (18,19), knee alignment (18,20), and features of osteoarthritis (21) in bones, but examples of measurement of soft tissue morphology in the knee are limited due to difficulties in segmenting soft tissues from CT images. In addition, new methods of assessing cartilage integrity using contrast agent diffusion are being developed (4,5,8,22-27). Reliable cartilage segmentation methods are required to ensure that areas of cartilage with increased attenuation due to contrast agent diffusion are included but that areas of similar attenuation where contrast agent has mixed with synovial fluid are excluded. Bone can be readily segmented from CT images without any contrast agent present (28), but cartilage, which can only be segmented *in vivo* in the presence of a contrast agent, is more difficult. Cartilage boundaries are clearest immediately after contrast injection. However, as contrast diffuses into the cartilage and with successive physiologic clearance of the contrast agent out of the joint, the sharp boundary between the cartilage and the surrounding fluid begins to blur over time. Clearance of the contrast agent out of the joint space reduces image contrast between the bone, iodine, and cartilage, making differentiation between tissues more challenging. If manual or threshold-based segmentation only identifies cartilage regions with minimal attenuation, the segmented cartilage morphology will be inaccurate due to exclusion of superficial tissue that contains diffused contrast agent. This is an important limitation as many studies assessing cartilage integrity using CT often use time-delayed images, or images taken up to 45 minutes after the initial injection (4,22). In addition, the ideal time between injection and CT imaging can vary between joints (29), is application-dependent, and cannot be predicted in advance due to patient variability, which could be a source of error in any cartilage morphology observations or calculation of other quantitative metrics following imaging.

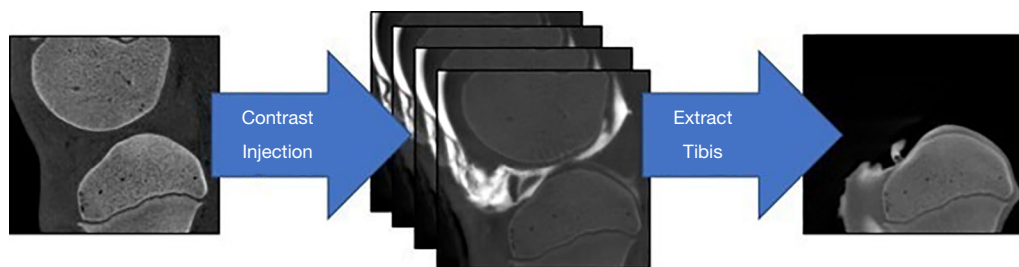
Previous studies have validated or used cartilage

segmentations from CT arthrography datasets, however these studies used manual or semi-automated segmentation methods which did not take diffusion into account and/or were not described in detail (4,23,24,30). Other similar methods (31) have been reported with promising results, but have not yet been validated against a ground truth measurement. In addition, segmentation methods are often validated using volume metrics (30) such as differences in total volume and Dice's coefficient, which are sufficient for computing quantitative parameters in a specific area of an image, but insufficient when trying to assess local cartilage surface morphology with precision over multiple scan time points. For potential applications such as tracking cartilage strain over time in CT arthrograms, detection of small changes in the cartilage surface location that do not have large effects on cartilage volume measurements is critical. As a result, this application requires a precise segmentation technique designed to detect the cartilage surface in a very specific area of the joint where there is contact between soft tissues.

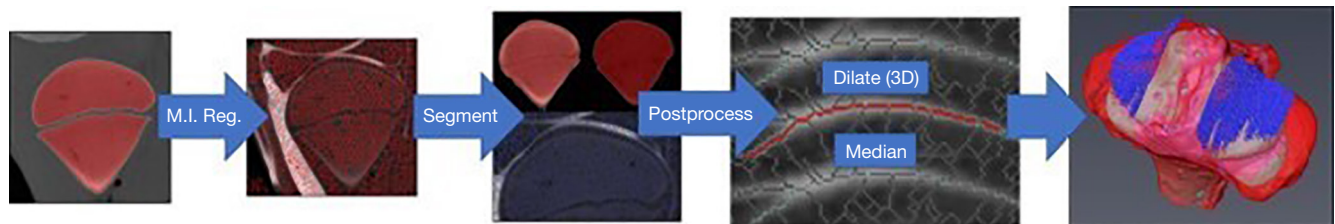
The goal of this study was to assess the utility of a watershed algorithm for the repeated measurement of cartilage surface location over time following contrast injection into bovine stifle joints. A watershed algorithm treats intensity of grayscale images as topology and picks up "basins", areas of continuous low attenuation, and "ridges", areas of local maximum attenuation. We consequently hypothesize that a watershed algorithm would be useful for our study as it picks up local maxima (ridges) in the datasets, and these maxima would be on the boundary of the cartilage surface. In areas where soft tissues are in contact with each other, such as cartilage-meniscus or cartilage-cartilage contact areas, there is no pool of highly concentrated contrast outside the tissues. Therefore, the surface layer of both tissues that contrast has diffused into is an area of local maximum attenuation.

## Methods

Three immature bovine stifle joints were purchased from a local butcher. Joints were first scanned using a Siemens Artis Zeego CT scanner without any contrast injected into the joint. Following the non-contrast scan, 40 mL of 50% iohexol contrast medium (350 mg iodine/mL), 50% phosphate buffered saline (PBS) was injected into the joint capsule. The joints were then manually articulated for 1 minute before being scanned in the supine position at 3, 8, 18, and 33 minutes after the end of the contrast injection.



**Figure 1** Image acquisition procedure. Images were taken first of the bovine stifle joint without contrast injection, then 3, 8, 18, and 33 minutes after injection. The tibia was then extracted and re-imaged.

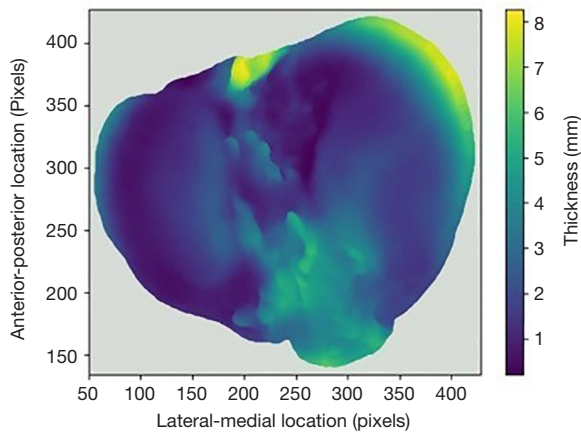


**Figure 2** Segmentation and image processing procedure. The initial in situ image taken before contrast injection was thresholded to segment the tibia. Watershed dam lines were applied to all in situ images taken after contrast injection in the sagittal plane. These images were segmented by following the applied watershed dam lines. Segmentations were dilated in 3D to fill any segmentation gaps where there were no watershed dam lines to follow. The median voxel in the axial direction was selected from the dilated initial segmentations to form the final segmentation. The image taken of the extracted tibia was segmented by thresholding to separate the tibia and soft tissues from the surrounding air.

Tibias were then extracted from the joints and scanned ex situ to capture the cartilage surface, a method shown to provide accurate morphology data (32). In this experiment, the surface layers of cartilage had contrast agent in them from the previous arthrographic scans, which increased contrast between the cartilage and the surrounding air and increased ease of segmentation. The experimental protocol is summarized in *Figure 1*. All scans were acquired in 19 seconds at 81 kVp with 496 projections. Images were reconstructed to a 0.2 mm isotropic voxel size using the scanner's reconstruction software.

The image processing procedure, described below, is summarized in *Figure 2*. All CT images were registered to the non-contrast in situ scan in Amira 6.3 (v6.3, FEI, Hillsboro, OR) using a mutual information algorithm. The images were then converted to 8-bit in Fiji (33). A 2D watershed algorithm (34,35) was applied to the in situ contrast scans in Fiji in the sagittal plane. The parameters of the watershed algorithm were: 3.0 pixel gaussian blur radius, dark objects/bright background, 4-connected, 0 minimum gray level, and 255 maximum gray level.

Dam lines were overlaid on to the images. Cartilage was segmented manually in Amira from all in situ contrast scans by selecting voxels along the dam lines that appeared to be on the cartilage surface in the sagittal plane for every slice of the images containing cartilage. During segmentation, a threshold was applied to the image so that only voxels on the watershed line could be selected and added to the segmentation. After all slices were segmented, the segmentations were dilated by 1 voxel in each direction in 3D and imported into MATLAB (r2016a, Mathworks, Natick, MA, USA) for postprocessing. The median voxel in the axial direction was selected for each scan if there was more than 1 voxel in the axial direction at any point on the transverse plane in the segmentation, with axial defined as along the axis of the reconstructed images closest to normal to the diaphysis of the tibia. If the median of the axial direction voxels was between two adjacent voxels, the most inferior voxel was chosen. This was done to repeatedly fill any 1–2 voxel wide gaps left in the segmentation where there were no watershed lines to follow, and to allow information from adjacent slices to inform segmentations of



**Figure 3** Representative cartilage thickness map based on a scan of the extracted tibia, superior view. Cartilage thickness ranged from less than 1 to 8 mm over the area of interest, which was the cartilage directly superior to the bone surface of the tibial plateau.

a given slice in the sagittal plane. The extracted tibia scans and the in situ scan without contrast were segmented in Amira by thresholding so that voxels including air and soft tissue, respectively, were not included in the segmentation. These segmentations are referred to as the reference segmentation and bone segmentation, respectively, throughout the paper. All segmentations were trimmed in MATLAB so that they did not extend past the tibial plateau surface in the transverse plane.

Binary segmentations of all cartilage and bone surfaces were analyzed using a custom Python 3 script. Surface point coordinates were extracted from binary segmentations. Cartilage thickness was calculated by finding the point on the bone surface closest to each cartilage point and taking the distance between the two points as the cartilage thickness. Differences in the Z (direction in the reconstruction closest to being aligned with the shaft of the femur) coordinate ( $\Delta z$ ) between watershed segmentations and the reference segmentation were calculated for each point on the transverse (XY) plane. Differences in thickness ( $\Delta t$ ) between the watershed segmentations and the reference segmentation were calculated by subtracting the cartilage thickness calculated at the corresponding point on the transverse plane of the reference segmentation from the thickness calculated at each point on the transverse plane of the watershed segmentation. Thickness maps, Z-coordinate difference ( $\Delta z$ ) maps, and thickness difference ( $\Delta t$ ) maps were calculated for all datasets. Histograms were calculated for  $\Delta z$  and  $\Delta t$  with 1-voxel (0.2 mm) wide bins.

Linear regression was used to analyze the relationship between cartilage thickness based on watershed segmentations and actual cartilage thickness. Dice's coefficients were calculated to compare the volume of watershed segmentations to the thresholded cartilage segmentation in areas where cartilage could be segmented using the watershed algorithm. This was done by excluding all voxels on the reference segmentation for which there was not a point at the same location on the transverse plane in the watershed segmentation from the calculation. Dice's coefficient (DSC) was calculated by applying Equation 1 to points in between and including the segmented cartilage surface and segmented bone surface for cartilage segmentation points that the reference and watershed segmentation had in common in the transverse plane:

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} \quad [1]$$

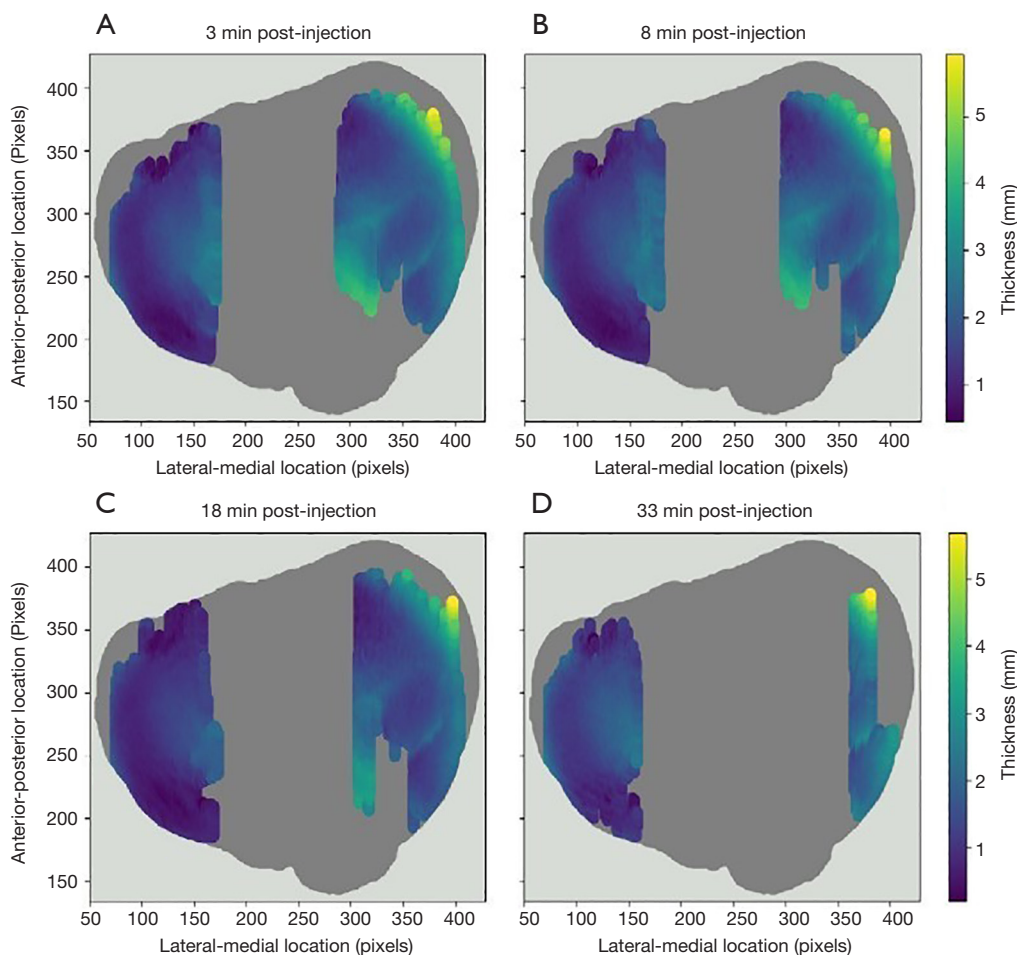
where X and Y are voxels between the cartilage surface and bone surface for the watershed and reference segmentations, respectively,  $|X \cap Y|$  represents the number of voxels the reference and watershed segmentation have in common, and  $|X|$  and  $|Y|$  represent the number of voxels in the watershed and reference segmentations, respectively.

## Results

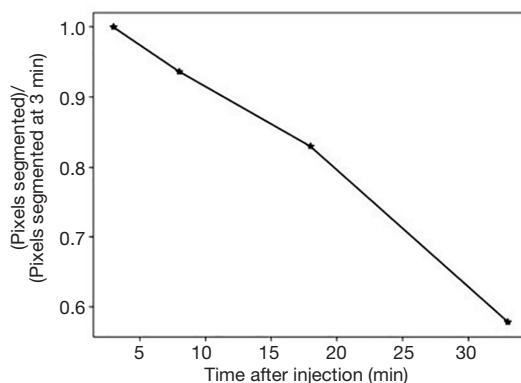
Cartilage thickness (*Figure 3*) in the area above the tibial plateau bone surface, which was the region of interest for this analysis, varied from less than 1mm in the thinnest areas to 8 mm in the thickest areas. Thickness maps (*Figure 4*) show that as time went on and the contrast agent redistributed and diffused in the joint space, fewer voxels could be segmented. This is quantified in *Figure 5*.

The ability to segment did not appear to depend on the cartilage thickness, but rather on the location in the joint space and the contrast agent distribution in the joint. This is evident from the observation that segmented regions shrank over time, but calculated cartilage thickness ranges stayed similar. When the contrast agent concentration dropped below the level where it was visible when images were converted to 8-bit per the watershed algorithm requirements, the watershed dam lines no longer reflected the cartilage surface location. In areas where this effect was obvious, affected voxels were not segmented. In areas where contrast concentration was low in the first scan, voxels were less likely to be segmented in scans taken at later time points.

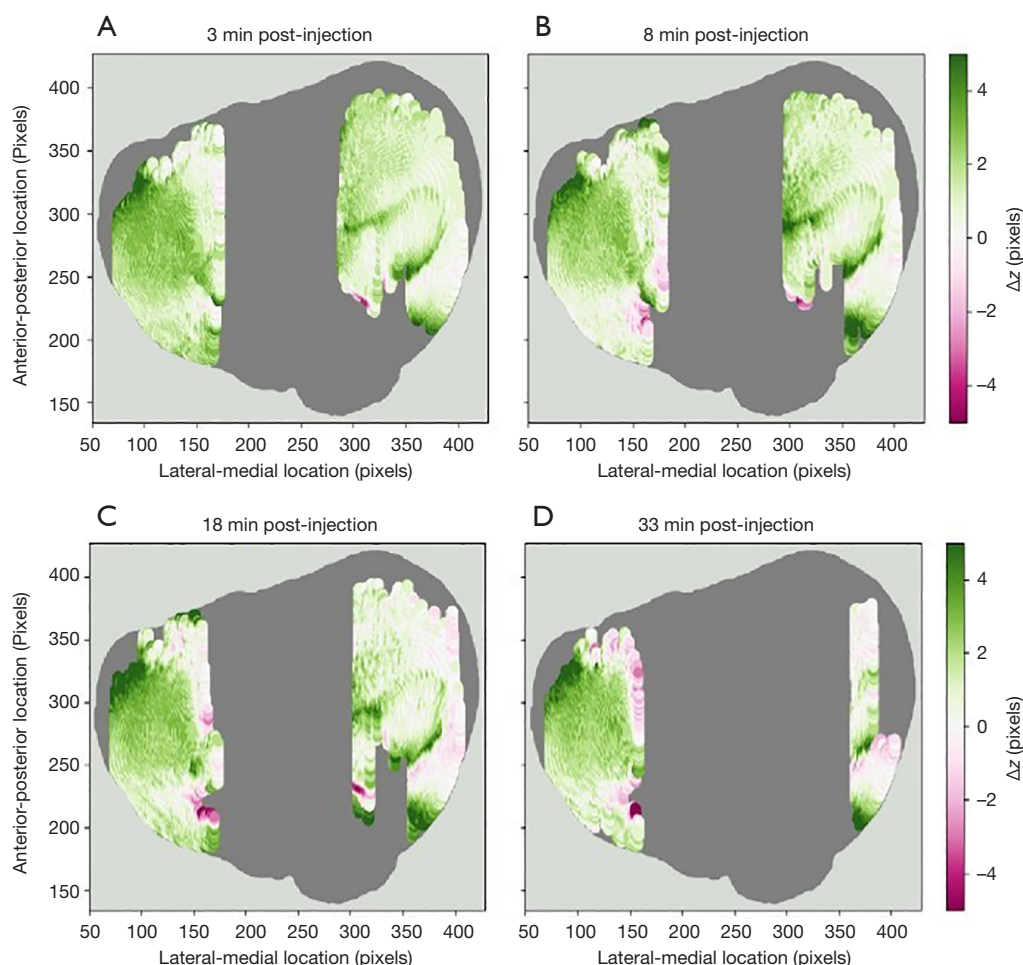
Voxels that could be segmented throughout all time



**Figure 4** Representative thickness maps for scans taken 3–33 minutes after iohexol injection, superior view. The entire cartilage surface could not be segmented at the 3-minute time point. The area that could be segmented using the watershed algorithm was less over time due to diffusion of the contrast agent into the joint tissues and the resulting concentration reduction in the joint space. General trends match the reference thickness, however there are identifiable areas with errors.



**Figure 5** Relationship between voxels segmented and time after injection. Numbers shown are for the sum of all scans of all 3 tibias. The number of voxels that could be segmented was lower at later time points because the contrast agent became more diffuse in the joint space and in the cartilage. When this occurred, the cartilage boundaries either blurred or did not have sufficient contrast to differentiate from the joint space.

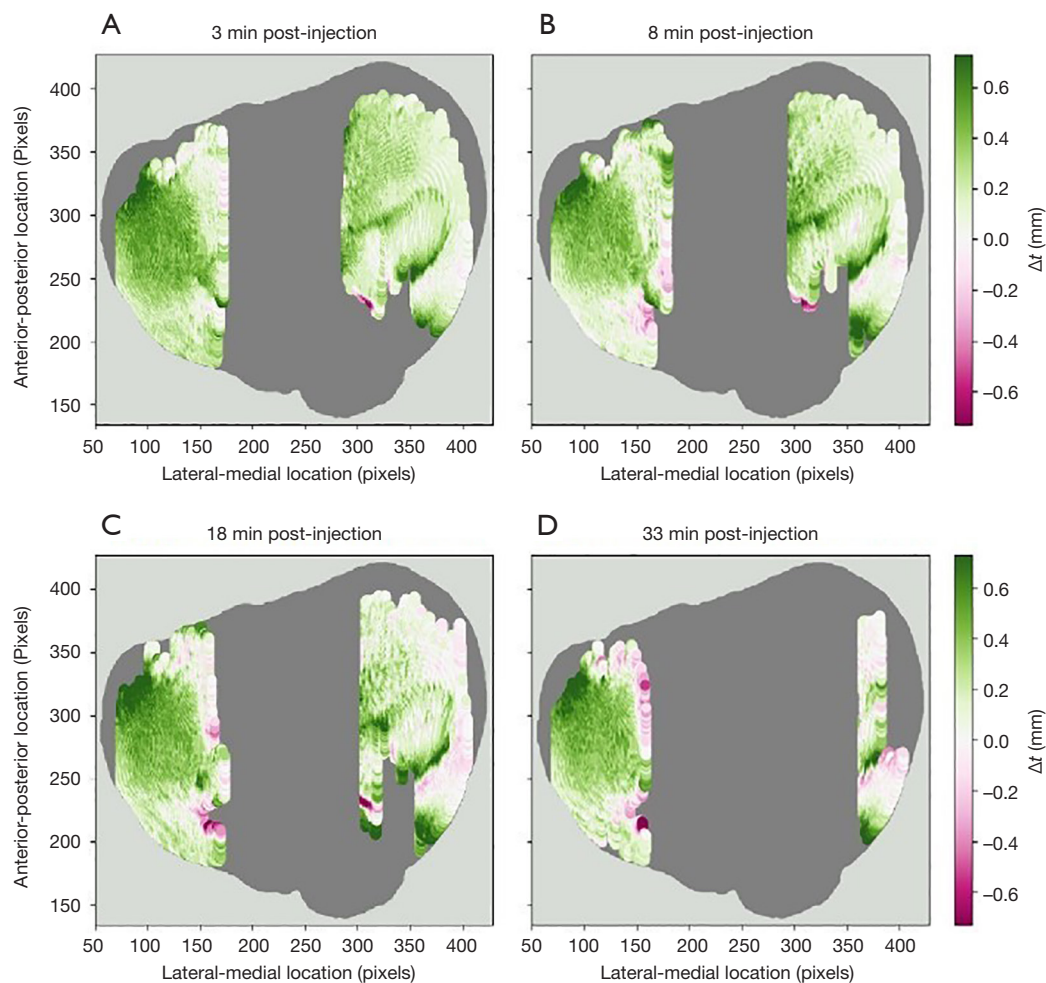


**Figure 6** Representative z-coordinate difference ( $\Delta z$ ) maps for scans taken 3–33 minutes after iohexol injection, superior view. Differences between the reference segmentation and watershed segmentation were mostly positive and up to 5 voxels in magnitude. The errors happened in discrete areas that are not related to cartilage thickness or anatomy.

points appeared to be close in z-location (*Figure 6*) and calculated thickness (*Figure 7*) over time, however there were variations of up to 0.8 mm in some discrete areas of the cartilage. The locations of these areas did not appear to be anatomy- or thickness-dependent. *Figure 8* shows examples of areas where segmentations matched or did not match actual thickness measurements. Segmentations were best in areas where there was a visible line of contrast between two tissues in direct contact, as seen in *Figure 8A*. Large pools of contrast, or areas where there was a relatively thick layer of contrast in between two tissues, caused overestimation of cartilage thickness because the local maximum attenuation was no longer at the cartilage surface, as seen in *Figure 8B, 8D*. Transition areas between

direct tissue contact, in this case cartilage and meniscus, and no direct contact also caused errors in thickness measurement, as seen in *Figure 8C*. There were areas where the watershed lines picked up a cartilage boundary in a relatively low attenuation area that could not be clearly segmented using an unguided manual segmentation. This is one of the advantages of using such an algorithm for guidance. Areas where errors occurred were consistent for all time points up to 33 minutes.

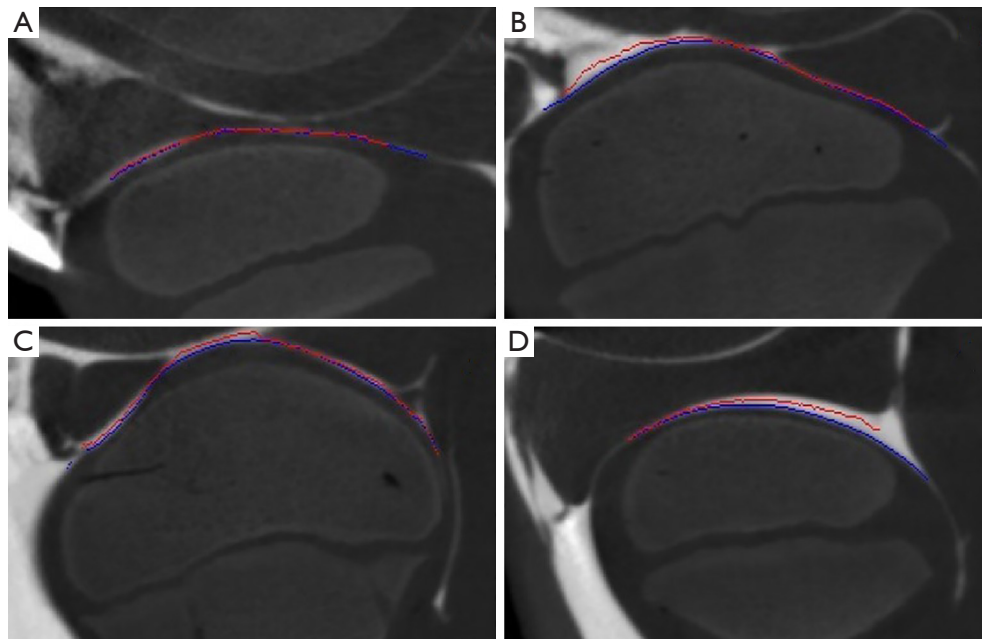
Overall, 60.8% of cartilage voxels segmented from the arthrograms were within 1 voxel in the axial direction of their location on the reference scan. 77.2% of cartilage thickness measurements made on the arthrogram scans were within 0.2 mm (1 voxel) of the thickness calculated



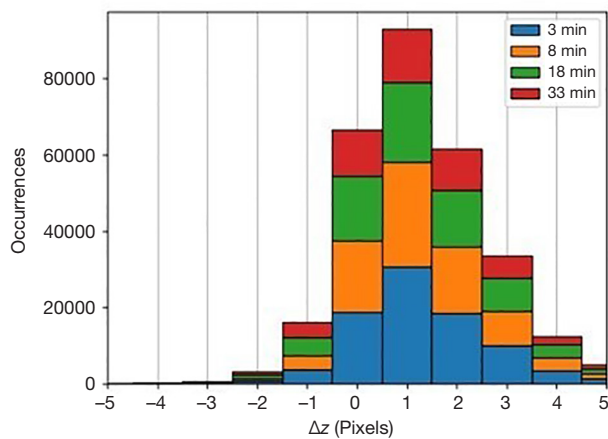
**Figure 7** Representative thickness difference ( $\Delta t$ ) maps for scans taken 3-33 minutes after iohexol injection, superior view. Because of the nearest neighbor method of calculating cartilage thickness, when the segmented voxel changed, the nearest point on the bone could also change, so there is not a 1:1 correlation between segmentation changes and thickness changes. Thickness differences were up to 0.7 mm in magnitude and, as with segmentation differences, occurred in discrete areas of the cartilage that were not related to anatomy or cartilage thickness.

for the corresponding point on the reference scan using the nearest neighbor thickness calculation method. The histogram in *Figure 9* shows a 1 voxel bias in the positive axial direction for the watershed segmentation. As a result, calculated cartilage thickness (*Figure 10*) was greater than actual cartilage thickness. Quantitative analysis shows that, over time, there was a slight decrease in the number of voxels segmented that were within 1 voxel of the reference segmentation (*Figure 11*). This resulted in a corresponding slight increase in  $\Delta t$  between the arthrogram segmentations and reference segmentations, as shown by slight widening of the distribution (*Figure 12*).

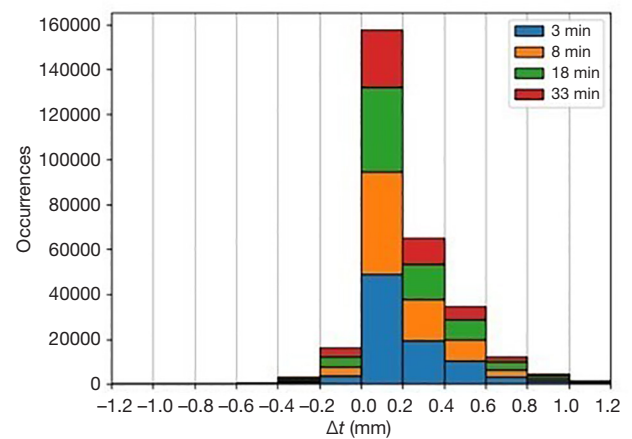
Linear regression analysis (*Figure 13*, *Table 1*) showed increased inaccuracy at low cartilage thicknesses (*Figure 14*). Correlation coefficients ranged from  $R^2=0.82-0.97$  for individual scans and decreased from  $R^2=0.93$  to  $R^2=0.88$  over the course of the 30 minutes between the first and last post injection scans. This is reflective of the widening error distribution over time observed in *Figure 12*. Overall, when points from all scan times and all tibias were included in the regression analysis,  $R^2=0.83$ . DSC for segmented areas of the arthrograms ranged from 0.91 to 0.95 (*Table 2*). There were no significant differences between DSC of dataset segmentations done at different time points.



**Figure 8** Examples of segmentation artifacts introduced by the watershed algorithm. Red voxels represent segmentations done at the 3-minute time point. Blue voxels represent the threshold-based segmentation of the extracted tibia. Image (A) is a good representation of when this segmentation method works, which is at moderate attenuation with clear contact between soft tissues. Images (B-D) show examples of areas where contrast can collect and cause a high attenuation area right outside the cartilage surface. This causes a plateau rather than a local maximum near the cartilage surface and leads to overestimation of thickness when the watershed algorithm is used for segmentation.

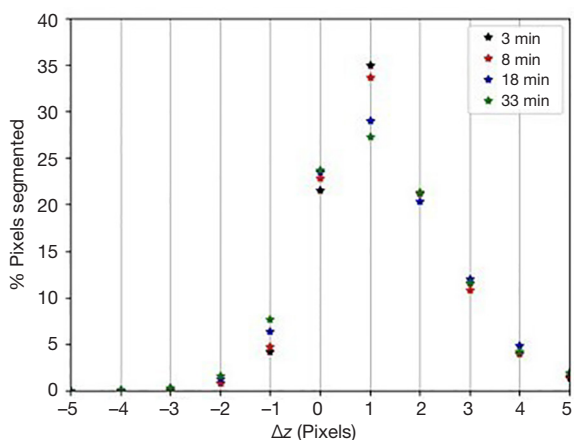


**Figure 9** Histogram for z-coordinate difference ( $\Delta z$ ) at each post-injection time point for all 3 tibias. Overall, there was a 1 voxel bias in the superior direction in the segmentations.

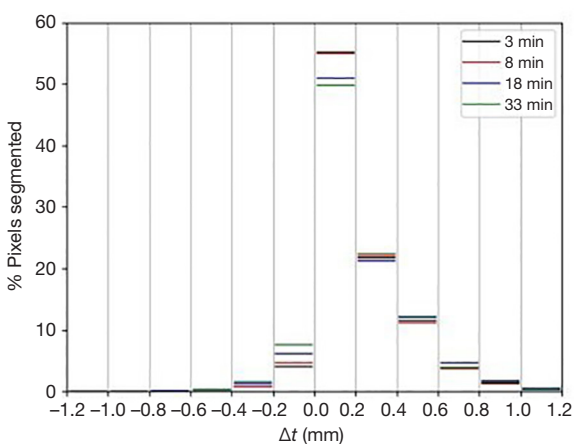


**Figure 10** Histogram for thickness difference ( $\Delta t$ ) at each post-injection time point for all 3 tibias. The calculated thickness bias due to segmentation error was predominantly in the range of +0–0.2 mm.





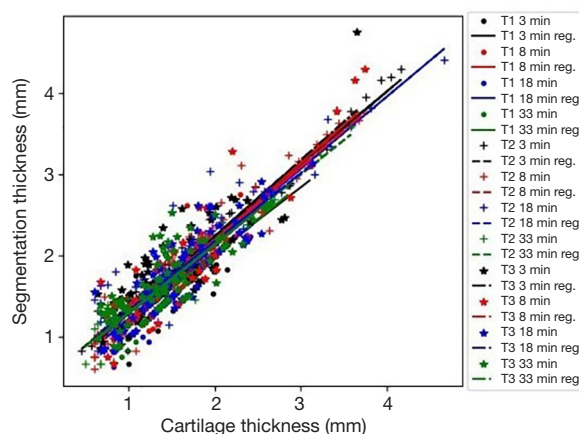
**Figure 11** Z-coordinate difference ( $\Delta z$ ) for each voxel on the transverse plane, shown for all tibias at each time point as a percentage. There was approximately a positive 1 voxel bias in the superior direction in segmentations. The percentage of voxels within 1 voxel of the reference segmentation went down over time.



**Figure 12** Thickness difference ( $\Delta t$ ), shown for all scans at each time point as a percentage. The percentage of voxels within 1 voxel length (0.2 mm) of the reference scan thickness fell slightly over time.

## Discussion

When a contrast solution is injected into a joint, it immediately starts diffusing into the soft tissues within the joint (36,37). This causes cartilage to appear thinner than it actually is when thresholding-based algorithms are used for segmentation (38), and the magnitude of this error is expected to increase with the time between injection and scanning. The watershed algorithm proved to be fairly



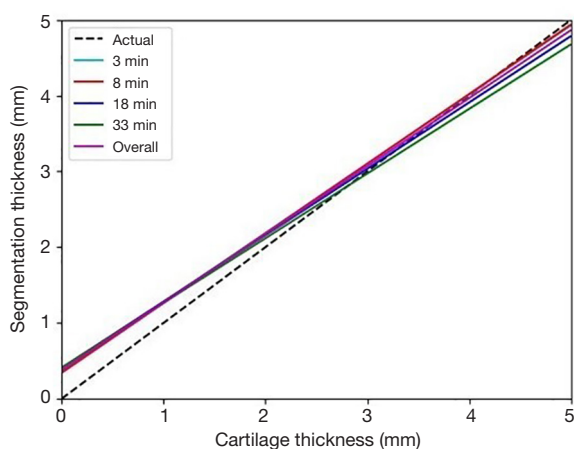
**Figure 13** Best fit linear regression plots for all scans of all tibias. There was good correlation between the reference cartilage thickness and thickness calculated using the watershed segmentations.

robust to this artifact, and despite visible blurring in the images over time, the segmentations using the watershed algorithm gave consistent thickness measurements in areas that could be segmented. This is important for contrast enhanced CT studies that involve analyzing cartilage morphology, and ensures that any observed morphology changes are due to actual changes in cartilage geometry, not the effects of diffusion or physiologic clearance of the contrast agent from the joint due to different post-injection scan times. This technique would also be ideal for tracking cartilage compression due to load bearing. In contrast to methods such as active contours and statistical models, a watershed segmentation approach does not require any assumptions or prior knowledge about the shape of the cartilage, and can therefore pick up irregularities such as compression patterns.

Though cartilage thickness measurements were consistent over time, the number of voxels that could be segmented went down as contrast diffused into a larger area and as a result became more diluted in the joint space. In practice, the limiting factor in comparing cartilage morphology across time points would be the last scan, where 40% fewer voxels could be segmented than at the first time point. In this study, due to the use of *ex vivo* bovine joints, the contrast agent diffused into surrounding joint tissues. *In vivo*, there is physiologic clearance in addition to diffusion, which may further reduce the contrast agent concentration,

**Table 1** Linear regression slopes (m), offsets in mm (b), and correlation coefficients ( $R^2$ ) for individual scans, each time point for all tibias, and overall for all scans. Correlation coefficients and the slope of the best fit lines dropped slightly over time, which could be a reflection of slightly reduced segmentation accuracy

Scan	Tibia 1			Tibia 2			Tibia 3			Overall		
	m	b	$R^2$	m	b	$R^2$	m	b	$R^2$	m	b	$R^2$
3 min	0.81	0.40	0.89	0.89	0.45	0.97	0.94	0.36	0.90	0.91	0.36	0.93
8 min	0.84	0.37	0.88	0.93	0.36	0.96	0.90	0.40	0.87	0.92	0.34	0.92
18 min	0.81	0.40	0.87	0.88	0.42	0.95	0.85	0.47	0.83	0.88	0.41	0.91
33 min	0.85	0.32	0.92	0.84	0.46	0.93	0.84	0.45	0.82	0.85	0.41	0.88
Overall										0.90	0.37	0.91



**Figure 14** Overall linear regression best fit lines for each post-injection scan time point. Overall, best fit lines at each time point were similar. The positive bias in thickness calculation shows up on this plot in the form of a positive offset. This is an indication that calculated thickness errors are not correlated with cartilage thickness.

and as a result the attenuation, image contrast, and number of voxels that can be segmented accurately as time elapses between injection and imaging. Therefore, though this technique could be used to make comparisons over time, the time scales it can be used for are somewhat limited.

Knowledge that local peak attenuation coincides with the cartilage surface could be leveraged to develop an automated cartilage surface detection algorithm. Such an algorithm would need to identify the cartilage surface as the local maximum in appropriate areas between input tibia and femur segmentations, such as areas of cartilage/cartilage or cartilage/meniscus contact, potentially adjust to an algorithm such as a half maximum (31) in areas where there

**Table 2** Dice's coefficient (DSC) calculated for segmented areas of each scan and averages for each time point. Though the DSCs are high, a 1 voxel bias could mean a difference of 4% to 30% thickness error depending on the area of the cartilage the segmentation is applied to. In addition, inclusion of non-cartilage voxels could adversely affect biomarkers calculated from computed tomography (CT) attenuation

Variable	3 min	8 min	18 min	33 min
Tibia 1	0.95	0.95	0.94	0.95
Tibia 2	0.94	0.94	0.94	0.92
Tibia 3	0.93	0.93	0.92	0.93
Mean	0.94	0.94	0.93	0.93
SD	0.01	0.01	0.01	0.01

is a contrast pool adjacent to the cartilage surface, and make use of machine learning methods for areas of uncertainty. Simulation of contrast agent diffusion into the cartilage may be helpful in determining the correct adjustments to make to locate cartilage boundaries in areas where there is contrast agent pooling as well. The segmentation method used in this study did not allow users to select voxels that were not on an automatically generated watershed line representing a local attenuation peak. Now that it has been verified that these local peak attenuation points coincide with the cartilage surface, an algorithm could be developed to detect peak attenuation points between the femur and tibia, and therefore the cartilage surface, automatically. This potential for automation is the power of the methods used in this study, and could enable an efficient image processing pipeline for tracking cartilage morphology over time in weight bearing scans.

For maximum achievable segmentation accuracy,

it would be ideal to modify experiments and clinical protocols to prevent pooling of the contrast agent in arthrograms used to assess cartilage morphology. Contrast agent injection volume and concentration, as well as time between injection and imaging, should be carefully chosen to balance minimizing errors related to diffusion, contrast agent clearance, and contrast agent pooling discussed above with being able to see accurate cartilage morphology in the largest area possible. In addition, a negatively charged ionic contrast agent could be helpful because it would not diffuse as much as the neutral one used in this experiment due to repulsion between the negative ions in the contrast agent and the negatively charged GAG molecules in the cartilage. Additional joint articulation could help to more evenly distribute contrast throughout the joint so that most of the cartilage can be segmented. The bovine legs used in this experiment do not articulate easily and have a limited range of motion, so spreading contrast through the joint if it was not distributed well after the original injection was a challenge. This affected contrast agent distribution, which in turn affected how many voxels could be segmented from the first scan using the watershed algorithm and how many fewer voxels could be segmented from later scans. For example, if there was very little contrast in an area of the joint in the first scan, then the contrast got fainter in later scans to the point where the watershed lines no longer resembled the cartilage surface. Likewise, if contrast collected in high concentration in an area during injection, and lack of joint mobility prevents it from redistributing during articulation, it would cause overestimation of cartilage thickness in some places and inability to segment due to the lack of a local maximum near the cartilage surface.

Linear regression analysis showed that segmentation errors were offset dominated, and as a result, larger where cartilage was thinner.  $R^2$  values were high and showed strong correlation between cartilage thickness measured from segmentations and validation scans. For all time points, the majority of thickness measurements were within 1 voxel of the validation scan. At the resolution of these datasets, this corresponds to a thickness error of 6–20% for cartilage 1–3 mm thick. Expected cartilage thickness of the population being scanned and the required accuracy and precision of the experiment should be taken into account when considering using arthrography to track cartilage morphology over time.

DSC of the cartilage segmentations did not differ significantly for all time points across scans. This could

be an indication that the cartilage segmentations are of sufficient quality to make comparisons across time. On the other hand, histograms showing segmentation error (*Figures 9,10*) indicate that this method does still have error in the surface regions. The images in *Figure 8* further support that, despite high DSC, segmentations can still deviate from actual surface geometry significantly. The discrepancy between DSC and accurate surface geometry increases with the volume of the segmented area. This must be considered when making image-based morphology measurements that require precise surface geometry. For applications such as cartilage strain measurement, the 1 voxel error in surface measurement could correspond to over 20% strain computation error in some areas of the cartilage. For diffusion measurements, inclusion of voxels that are mostly fluid would likely increase the calculated diffusion coefficient of the cartilage. For these applications, average surface distance could be a better choice than DSC for evaluating overall segmentation accuracy, and histograms could be useful to establish how many voxels are within a specific required accuracy range.

Segmentation errors happen primarily because of contrast agent collecting in areas directly adjacent to the cartilage surface. In areas where tissues were in direct contact, segmentations were accurate because the local maximum corresponded to the small, contrast agent containing fluid layer between the tissues, which was also the cartilage surface. In areas where there was not direct contact between tissues, the cartilage surface was no longer the local maximum. Rather, there was a plateau in attenuation outside the cartilage surface due to a contrast pool. This caused overestimation of cartilage thickness in those areas.

Multiple studies have shown that arthrograms or even the diffusion artifact itself have potential to be imaging biomarkers for early signs of cartilage degradation (39–42). Being able to reliably segment images that include some diffusion of contrast into the joint space could enable the use of these biomarkers *in vivo*. If the cartilage surface cannot be identified, the amount of contrast agent diffusing in the cartilage will not be quantified correctly. This, in turn, will lead to inaccurate diffusivity or contrast agent partition calculations. Segmentation methods such as the one discussed here can help ensure that contrast diffusion does not change the apparent surface location in the image. If used with a registration algorithm to prevent any artifacts related to patient motion, the cartilage surface can be reliably identified across scans and metrics that require multiple scan time points, such as diffusivity or strain rates

under load, can be quantified accurately.

Though MRI has been shown to be more sensitive in lesion detection than CT arthrography (43), CT arthrography is still used as a primary diagnostic in situations where cost, medical device implants, tattoos, time constraints, or various other circumstances make it a better choice. In addition, delayed protocols can provide comparable information to contrast enhanced MRI, which has been shown to provide information about cartilage integrity (22), and correlates with EPIC- $\mu$ CT, a common method used to assess cartilage in small animal models (24). Segmentation methods such as the one described above, when automated, have the potential to make analysis of CT data more robust, especially given the changing contrast agent distribution in the joint over time. This could lead to better CT-based biomarkers, and a wider variety of diagnostics to assess patients' cartilage health and detect potential diseases early on.

## Conclusions

Following watershed dam lines shows promise as a technique for precisely identifying the cartilage surface in areas where soft tissues are in contact with each other for contrast enhanced CT data in the bovine stifle. This technique could enable detection of small changes in the cartilage surface, and measurements such as compressive strain, over multiple scan time points. Pooling of contrast agent does cause predictable inaccuracies, and segmentations could potentially be done in a way that corrects for it to decrease error. This method, because it is based on an automated algorithm, has the potential to be automated in the future by developing an algorithm which locates the area of interest based on femur and tibia segmentations, and picks the correct peak attenuation point between the two bones. This could be informed by experimentation on contrast agent diffusion patterns in cartilage. Experimental parameters such as contrast agent concentration, volume of contrast injected, and articulation time may be able to be changed to enable better segmentation using this method as well.

## Acknowledgments

*Funding:* This work was supported by the National Institute of Arthritis, Musculoskeletal and Skin Disorders (NIAMS) of the National Institutes of Health under award number R01AR065248 and by a Stanford Bio-X Graduate Fellowship to MEH.

## Footnote

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/qims-20-1062>). GEG serves as an unpaid editorial board member of *Quantitative Imaging in Medicine and Surgery*. MEH reports that this research was supported by grant # R01AR065248 from the NIAMS of the U.S. National Institutes of Health and a graduate student fellowship from Stanford Bio-X. MSB reports that Natural Sciences and Engineering Research Council of Canada (NSERC) graduate student fellowship received for duration of work conducted. GEG reports that this research was supported by grant # R01AR065248 from the NIAMS of the U.S. National Institutes of Health. MEL reports that this research was supported by grant # R01AR065248 from the NIAMS of the U.S. National Institutes of Health.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Bitton R. The economic burden of osteoarthritis. *Am J Manag Care* 2009;15:S230-5.
2. Allen BC, Peters CL, Brown NA, Anderson AE. Acetabular cartilage thickness: accuracy of three-dimensional reconstructions from multidetector CT arthrograms in a cadaver study. *Radiology* 2010;255:544-52.
3. Xie L, Lin AS, Levenston ME, Guldberg RE. Quantitative assessment of articular cartilage morphology via EPIC-microCT. *Osteoarthritis Cartilage* 2009;17:313-20.
4. Kokkonen HT, Suomalainen JS, Joukainen A, Kröger H, Sirola J, Jurvelin JS, Salo J, Töyräs J. In vivo diagnostics of human knee cartilage lesions using delayed CBCT arthrography. *J Orthop Res* 2014;32:403-12.

5. Piscaer TM, Waarsing JH, Kops N, Pavljasevic P, Verhaar JA, van Osch GJ, Weinans H. In vivo imaging of cartilage degeneration using microCT-arthrography. *Osteoarthritis Cartilage* 2008;16:1011-7.
6. Kotwal N, Li J, Sandy J, Plaas A, Sumner DR. Initial application of EPIC- $\mu$ CT to assess mouse articular cartilage morphology and composition: effects of aging and treadmill running. *Osteoarthritis Cartilage* 2012;20:887-95.
7. Palmer AW, Guldborg RE, Levenston ME. Analysis of cartilage matrix fixed charge density and three-dimensional morphology via contrast-enhanced microcomputed tomography. *Proc Natl Acad Sci U S A* 2006;103:19255-60.
8. Siebelt M, van Tiel J, Waarsing JH, Piscaer TM, van Straten M, Booij R, Dijkshoorn ML, Kleinrensink GJ, Verhaar JA, Krestin GP, Weinans H, Oei EH. Clinically applied CT arthrography to measure the sulphated glycosaminoglycan content of cartilage. *Osteoarthritis Cartilage* 2011;19:1183-9.
9. Tuominen EK, Kankare J, Koskinen SK, Mattila KT. Weight-bearing CT imaging of the lower extremity. *AJR Am J Roentgenol* 2013;200:146-8.
10. Kang DH, Kang C, Hwang DS, Song JH, Song SH. The value of axial loading three dimensional (3D) CT as a substitute for full weightbearing (standing) 3D CT: Comparison of reproducibility according to degree of load. *Foot Ankle Surg* 2019;25:215-20.
11. Choi JH, Maier A, Keil A, Pal S, McWalter EJ, Beaupré GS, Gold GE, Fahrig R. Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. II. Experiment. *Med Phys* 2014;41:061902.
12. Choi JH, Fahrig R, Keil A, Besier TF, Pal S, McWalter EJ, Beaupré GS, Maier A. Fiducial marker-based correction for involuntary motion in weight-bearing C-arm CT scanning of knees. Part I. Numerical model-based optimization. *Med Phys* 2013;40:091905.
13. Berger M, Müller K, Aichert A, Unberath M, Thies J, Choi JH, Fahrig R, Maier A. Marker-free motion correction in weight-bearing cone-beam CT of the knee joint. *Med Phys* 2016;43:1235-48.
14. Unberath M, Choi JH, Berger M, Maier A, Fahrig R. Image-based compensation for involuntary motion in weight-bearing C-arm cone-beam CT scanning of knees. In *Medical Imaging 2015: Image Processing* 2015;9413:94130D. International Society for Optics and Photonics.
15. Syben C, Bier B, Berger M, Aichert A, Fahrig R, Gold G, Levenston M, Maier A. Joint calibration and motion estimation in weight-bearing cone-beam CT of the knee joint using fiducial markers. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017) 2017:494-7.
16. Bier B, Ravikumar N, Unberath M, Levenston M, Gold G, Fahrig R, Maier A. Range imaging for motion compensation in C-arm cone-beam CT of knees under weight-bearing conditions. *Journal of Imaging* 2018;4:13.
17. Bier B, Aschoff K, Syben C, Unberath M, Levenston M, Gold G, Fahrig R, Maier A. Detecting anatomical landmarks for motion estimation in weight-bearing imaging of knees. In *International Workshop on Machine Learning for Medical Image Reconstruction*. 2018:83-90. Springer, Cham.
18. Hirschmann A, Buck FM, Fucntese SF, Pfirrmann CW. Upright CT of the knee: the effect of weight-bearing on joint alignment. *Eur Radiol* 2015;25:3398-404.
19. Thawait GK, Demehri S, AlMuhit A, Zbijweski W, Yorkston J, Del Grande F, Zikria B, Carrino JA, Siewerdsen JH. Extremity cone-beam CT for evaluation of medial tibiofemoral osteoarthritis: Initial experience in imaging of the weight-bearing and non-weight-bearing knee. *Eur J Radiol* 2015;84:2564-70.
20. Hirschmann A, Buck FM, Herschel R, Pfirrmann CWA, Fucntese SF. Upright weight-bearing CT of the knee during flexion: changes of the patellofemoral and tibiofemoral articulations between 0° and 120°. *Knee Surg Sports Traumatol Arthrosc* 2017;25:853-62.
21. Segal NA, Nevitt MC, Lynch JA, Niu J, Torner JC, Guermazi A. Diagnostic performance of 3D standing CT imaging for detection of knee osteoarthritis features. *Phys Sportsmed* 2015;43:213-20.
22. Hirvasniemi J, Kulmala KA, Lammentausta E, Ojala R, Lehenkari P, Kamel A, Jurvelin JS, Töyräs J, Nieminen MT, Saarakkala S. In vivo comparison of delayed gadolinium-enhanced MRI of cartilage and delayed quantitative CT arthrography in imaging of articular cartilage. *Osteoarthritis Cartilage* 2013;21:434-42.
23. Kokkonen HT, Aula AS, Kröger H, Suomalainen JS, Lammentausta E, Mervaala E, Jurvelin JS, Töyräs J. Delayed Computed Tomography Arthrography of Human Knee Cartilage In Vivo. *Cartilage* 2012;3:334-41.
24. van Tiel J, Siebelt M, Reijman M, Bos PK, Waarsing JH, Zuurmond AM, Nasserinejad K, van Osch GJ, Verhaar JA, Krestin GP, Weinans H, Oei EH. Quantitative in vivo CT arthrography of the human osteoarthritic knee to estimate cartilage sulphated glycosaminoglycan content: correlation with ex-vivo reference standards. *Osteoarthritis Cartilage* 2016;24:1012-20.

25. Stewart RC, Honkanen JTJ, Kokkonen HT, Tiitu V, Saarakkala S, Joukainen A, Snyder BD, Jurvelin JS, Grinstaff MW, Töyräs J. Contrast-Enhanced Computed Tomography Enables Quantitative Evaluation of Tissue Properties at Intra-joint Regions in Cadaveric Knee Cartilage. *Cartilage* 2017;8:391-9.
26. Bansal PN, Joshi NS, Entezari V, Malone BC, Stewart RC, Snyder BD, Grinstaff MW. Cationic contrast agents improve quantification of glycosaminoglycan (GAG) content by contrast enhanced CT imaging of cartilage. *J Orthop Res* 2011;29:704-9.
27. Myller KA, Turunen MJ, Honkanen JT, Väänänen SP, Iivarinen JT, Salo J, Jurvelin JS, Töyräs J. In Vivo Contrast-Enhanced Cone Beam CT Provides Quantitative Information on Articular Cartilage and Subchondral Bone. *Ann Biomed Eng* 2017;45:811-8.
28. Waarsing JH, Day JS, Weinans H. An improved segmentation method for in vivo microCT imaging. *J Bone Miner Res* 2004;19:1640-50.
29. Omoumi P, Mercier GA, Lecouvet F, Simoni P, Vande Berg BC. CT arthrography, MR arthrography, PET, and scintigraphy in osteoarthritis. *Radiol Clin North Am* 2009;47:595-615.
30. Haubner M, Eckstein F, Schnier M, Löscher A, Sittek H, Becker C, Kolem H, Reiser M, Englmeier KH. A non-invasive technique for 3-dimensional assessment of articular cartilage thickness based on MRI. Part 2: Validation using CT arthrography. *Magn Reson Imaging* 1997;15:805-13.
31. Myller KAH, Honkanen JTJ, Jurvelin JS, Saarakkala S, Töyräs J, Väänänen SP. Method for Segmentation of Knee Articular Cartilages Based on Contrast-Enhanced CT Images. *Ann Biomed Eng* 2018;46:1756-67.
32. Akiyama K, Sakai T, Koyanagi J, Murase T, Yoshikawa H, Sugamoto K. Three-dimensional distribution of articular cartilage thickness in the elderly cadaveric acetabulum: a new method using three-dimensional digitizer and CT. *Osteoarthritis Cartilage* 2010;18:795-802.
33. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;9:676-82.
34. Tsukahara M, Mitrovic S, Gajdosik V, Margaritondo G, Pournin L, Ramaioli M, Sage D, Hwu Y, Unser M, Liebling TM. Coupled tomography and distinct-element-method approach to exploring the granular media microstructure in a jamming hourglass. *Phys Rev E Stat Nonlin Soft Matter Phys* 2008;77:061306.
35. Sage D, Unser M. Teaching image-processing programming in Java. *IEEE Signal Processing Magazine* 2003;20:43-52.
36. Silvast TS, Kokkonen HT, Jurvelin JS, Quinn TM, Nieminen MT, Töyräs J. Diffusion and near-equilibrium distribution of MRI and CT contrast agents in articular cartilage. *Phys Med Biol* 2009;54:6823-36.
37. Kulmala KA, Korhonen RK, Julkunen P, Jurvelin JS, Quinn TM, Kröger H, Töyräs J. Diffusion coefficients of articular cartilage for different CT and MRI contrast agents. *Med Eng Phys* 2010;32:878-82.
38. Anderson AE, Ellis BJ, Peters CL, Weiss JA. Cartilage thickness: factors influencing multidetector CT measurements in a phantom study. *Radiology* 2008;246:133-41.
39. Siebelt M, Waarsing JH, Kops N, Piscoer TM, Verhaar JA, Oei EH, Weinans H. Quantifying osteoarthritic cartilage changes accurately using in vivo microCT arthrography in three etiologically distinct rat models. *J Orthop Res* 2011;29:1788-94.
40. Thote T, Lin AS, Raji Y, Moran S, Stevens HY, Hart M, Kamath RV, Guldborg RE, Willett NJ. Localized 3D analysis of cartilage composition and morphology in small animal models of joint degeneration. *Osteoarthritis Cartilage* 2013;21:1132-41.
41. Tamura S, Nishii T, Shiomi T, Yamazaki Y, Murase K, Yoshikawa H, Sugano N. Three-dimensional patterns of early acetabular cartilage damage in hip dysplasia; a high-resolution CT arthrography study. *Osteoarthritis Cartilage* 2012;20:646-52.
42. Bansal PN, Stewart RC, Entezari V, Snyder BD, Grinstaff MW. Contrast agent electrostatic attraction rather than repulsion to glycosaminoglycans affords a greater contrast uptake ratio and improved quantitative CT imaging in cartilage. *Osteoarthritis Cartilage* 2011;19:970-6.
43. Rand T, Brossmann J, Pedowitz R, Ahn JM, Haghigi P, Resnick D. Analysis of patellar cartilage. Comparison of conventional MR imaging and MR and CT arthrography in cadavers. *Acta Radiol* 2000;41:492-7.

**Cite this article as:** Hall ME, Black MS, Gold GE, Levenston ME. Validation of watershed-based segmentation of the cartilage surface from sequential CT arthrography scans. *Quant Imaging Med Surg* 2022;12(1):1-14. doi: 10.21037/qims-20-1062