Open Source Software for Automatic Subregional Assessment of Knee Cartilage Degradation Using Quantitative T2 Relaxometry and Deep Learning

Kevin A. Thomas1, Dominik Krzemiński, MS2, Łukasz Kidziński, PhD3, Rohan Paul, MS1, Elka B. Rubin4, Eni Halilaj, PhD5, Marianne S. Black, PhD4, Akshay Chaudhari, PhD1,4, Garry E. Gold, MD3,4,6, and Scott L. Delp, PhD3,6,7

Abstract
Objective. We evaluated a fully automated femoral cartilage segmentation model for measuring T2 relaxation values and longitudinal changes using multi-echo spin-echo (MESE) magnetic resonance imaging (MRI). We open sourced this model and developed a web app available at https://kl.stanford.edu into which users can drag and drop images to segment them automatically. Design. We trained a neural network to segment femoral cartilage from MESE MRIs. Cartilage was divided into 12 subregions along medial-lateral, superficial-deep, and anterior-central-posterior boundaries. Subregional T2 values and four-year changes were calculated using a radiologist’s segmentations (Reader 1) and the model’s segmentations. These were compared using 28 held-out images. A subset of 14 images were also evaluated by a second expert (Reader 2) for comparison. Results. Model segmentations agreed with Reader 1 segmentations with a Dice score of 0.85 ± 0.03. The model’s estimated T2 values for individual subregions agreed with those of Reader 1 with an average Spearman correlation of 0.89 and average mean absolute error (MAE) of 1.34 ms. The model’s estimated four-year change in T2 for individual subregions agreed with Reader 1 with an average correlation of 0.80 and average MAE of 1.72 ms. The model agreed with Reader 1 at least as closely as Reader 2 agreed with Reader 1 in terms of Dice score (0.85 vs. 0.75) and subregional T2 values. Conclusions. Assessments of cartilage health using our fully automated segmentation model agreed with those of an expert as closely as experts agreed with one another. This has the potential to accelerate osteoarthritis research.

Keywords
osteoarthritis, cartilage segmentation, T2 map MRI, neural network, deep learning

Introduction
Knee osteoarthritis (OA) affects 20-40% of the U.S. population older than 65 years and currently has no cure.1,3 Diagnosis and measurement of patients’ OA severity typically relies on the use of X-rays, which only enable the detection and assessment of OA that has progressed to the point of joint space narrowing and visible changes to the bone.4,5 Novel quantitative magnetic resonance imaging (qMRI) techniques have the potential to measure early changes in cartilage matrix composition before the onset of gross structural changes.6,7 For example, T2 relaxation time mapping from multi-echo spin-echo (MESE) T2-weighted MRIs potentially reflects the hydration level and collagen of the matrix.8 As OA progresses, T2 values in affected cartilage tend to increase.9 Because the disease takes years to progress at the structural scale, T2 relaxometry potentially offers the ability to measure the effect of therapeutic
interventions earlier than is possible with radiographs or structural MRI. Early measurement of cartilage changes may also enable earlier intervention.

Currently, manual or semi-automated segmentation of cartilage in MRI scans is a necessary first step for assessing cartilage health from these images. Manual segmentation is time-intensive. This bottleneck leads researchers to restrict their analyses to small subregions of cartilage or individual MRI slices, losing valuable information. Most work on automatic segmentation of cartilage in MRIs has focused on structural sequences used for morphological assessments (e.g., thickness and volume). While these sequences feature higher cartilage-to-background contrast and higher resolution than qMRI methods, the segmentation of MESE MRIs is crucial for the assessment of early changes in cartilage matrix composition. Prior studies have used image registration to transfer the segmentation of morphological images to corresponding MESE images of the same patient taken in succession. However, registration can be imperfect, partially due to the potential for non-affine movements of the knee throughout the acquisition time of the two images (e.g., flexion or extension of the joint or compression of soft tissues). The risk of patient movement and registration error may be higher when the time between the acquisition of the morphological image and T2 image is longer. For example, the image acquisition protocol for the Osteoarthritis Initiative (OAI), a longitudinal study of OA, separated acquisition of the morphological images from the MESE images by 18 minutes. Differences in the contrast and resolution between the morphological images and MESE images also contribute to imperfect registration.

A few recent studies have aimed to automate the segmentation of MESE images directly. These have used atlas-based approaches or a combination of deep learning with simplex deformable modeling. While they have reported promising results, the ability of fully automated methods to produce accurate, subregional measurements of cross-sectional and longitudinal variations in T2 has not been explored. The aim of this work was to develop a fully automated femoral cartilage segmentation model that operates directly on MESE images and to evaluate the model as a means to measure individuals’ subregional T2 values longitudinally. A model that can take a MESE MRI as input and produce expert-quality assessments of T2-based cartilage health and disease progression would enable more accurate, efficient OA research.

**Methods**

**Overview**

We used images from OAI subjects. The OAI study was institutional review board (IRB) compliant. All subjects provided written informed consent. Our analysis was approved by our institution’s IRB (eProtocol #31592). Each MRI was segmented by a musculoskeletal radiologist. These segmentations were used as ground truth. We used a convolutional neural network (CNN) model to learn MRI features predictive of cartilage location. Agreement between predicted segmentations and ground-truth segmentations was assessed in a held-out test set of subjects. Segmented cartilage was divided into 12 subregions. The average T2 value and four-year change in T2 value was calculated for each subregion of each test set subject using the radiologist’s segmentations and the CNN’s segmentation. Focal areas of increased T2 value were identified automatically and the percentage of total cartilage area covered by these lesions was compared between segmentation approaches.

**Data**

The OAI is a public study of knee OA in which MRIs were collected longitudinally. We used 286 sagittal plane MESE MRI volumes (Table 1) from 143 OAI subjects assessed at baseline and four years later. Half of the subjects were randomly selected from among those in the OAI Incidence Cohort with body mass index (BMI) >30 kg/m2 and the other half were age- and sex-matched controls with normal BMI in the Incidence Cohort. All had a Kellgren-Lawrence grade of 0 in the imaged knee, indicating no radiographic OA. All were determined to be at risk for developing OA as determined by knee symptoms and at least two other risk factors (e.g., family history, previous knee injury, occupational burden). At-risk subjects with no radiographic OA were selected in order to assess the model’s ability to enable assessment of OA before it is visible on X-ray. This is an important use case for T2 map MRI. Subjects were 48% female, 90% Caucasian/9% Black or African American/1% Asian, aged 45 to 78 years, and had BMIs of 18.4 to 44.3 kg/m2. Each image was segmented with a semi-automated process and refined by a musculoskeletal radiologist with 15 years of experience, referred to as Reader 1. Subjects were split into training (115 subjects, 138 males, 168 females) and validation sets (68 subjects, 86 males, 106 females) with equal gender distribution.

**Table 1. Osteoarthritis Initiative Multi-Echo Spin-Echo Magnetic Resonance Imaging Parameters.**

| Parameter                        | Value |
|----------------------------------|-------|
| Matrix (phase)                   | 269   |
| Matrix (frequency)               | 384   |
| Number of slices                 | 21    |
| Field of view (mm)               | 120   |
| Slice thickness/gap (mm/mm)      | 3 / 0.5 |
| Echo times (ms)                  | 10, 20, 30, 40, 50, 60, 70 |
| Repetition time (ms)             | 2700  |
| X-resolution (mm)                | 0.313 |
| Y-resolution (mm)                | 0.446 |
230 image volumes), validation (14 subjects, 28 image volumes), and test (14 subjects, 28 image volumes) sets with no crossover of subjects across sets. The subjects’ OAI patient identification numbers and train/validation/test set assignments are available in Supplementary Material 1.

**Segmentation Model Training**

A model’s architecture determines the strategy for how it maps an input (e.g., MRI slice) to an output (e.g., segmentation mask). We used the two-dimensional U-Net, a CNN architecture shown to perform well on medical image segmentation tasks, including cartilage segmentation in morphological MRIs. We trained the CNN to learn MRI features that were predictive of cartilage location from the dataset. It was optimized to produce segmentations with high Dice scores relative to the ground truth segmentations (Supplementary Material 2: Loss Function). The architecture was designed to segment each slice independently.

We trained several models, all with the U-Net architecture, but different hyperparameters (Supplementary Material 3: Training Hyperparameters). Each model was evaluated with the validation set images. The model with the highest average validation set Dice score was selected as the final model and was then evaluated with the test set. This enabled the final model’s performance on the test set to serve as an objective measure of how it would perform on new images not used in this study. See Supplementary Material 4 for specifications of the hardware and software used.

**Image Preprocessing**

The model was designed to take in the second echo of a slice to produce that slice’s segmentation. However, while training the network, a given slice’s first echo was used with 20% probability, the second echo was used with 60% probability, and the third echo was used with 20% probability. This echo selection process was performed for each slice in each epoch (i.e., each round of training in which a model sees every training image once). Exposing the model to a range of decay times was done to improve the model’s performance on new patients. To increase the model’s sensitivity, 90% of slices that did not contain cartilage were randomly removed from each image volume in each epoch. Only the second echo was used when evaluating models with the validation and test sets and all slices of these images were retained.

Before an echo image was used as input to the model, its voxel values were rescaled so its median voxel value was 0, its 25th percentile voxel value was −1, and its 75th percentile voxel value was 1. Voxel values were trimmed between 3rd and 97th percentile to remove outliers.

**Segmentation Refinement**

For each voxel of an input image slice, the CNN outputs a value, \( P \in [0,1] \). Larger values in this range can be interpreted as voxels that the model predicts to contain cartilage. All voxels with \( P > 0.01 \) were initially considered as potentially containing cartilage. T2 values were calculated for each of these voxels for both expert and model segmentations. The first echo was excluded to minimize stimulated-echo artifacts and a noise-corrected monoeponential fit was used. Segmented voxels with a T2 value outside the physiological range of cartilage, \([0, 100] \) ms, were discarded. This refinement procedure was done for both ground truth segmentations and the model’s segmentations.

In early experiments, models frequently predicted small amounts of cartilage at the medial and lateral joint margins in slices that did not contain cartilage. A threshold was therefore used to set a minimum number of cartilage voxels per slice. In slices that were predicted to have fewer cartilage voxels than the threshold, all of these voxels were discarded from the model’s segmentation. Another threshold was applied to \( P \) such that any \( P \geq \) threshold was considered cartilage and all other outputs were not. The validation image set was used to set the values of these two thresholds to maximize the Dice score between model segmentations and Reader 1 segmentations. These thresholds, identified to be 425 voxels and \( P \geq 0.501 \) via grid search, were applied to all test set segmentations to binarize them.

**Segmentation Evaluation**

**Direct Comparison on Segmentation Masks.** To compare the agreement between the model and Reader 1, the volumetric Dice score and Jaccard index were calculated for each image’s segmentations. To understand how this compares with inter-reader agreement, an image volume from each test set subject, 14 image volumes total, was manually segmented by a researcher with 3 years of cartilage segmentation experience, referred to as Reader 2, for comparison with the ground truth segmentations. These segmentations were reviewed by a musculoskeletal radiologist with 21 years of experience who had previously served on the Imaging Advisory Board for the OAI.

**Comparison of Subregional T2 Mean Values.** To investigate the model’s impact on T2 relaxometry, we used it to evaluate the average T2 value and four-year change in average T2 value in anatomical subregions of the femoral cartilage delineated along medial-lateral, superficial-deep, and anterior-central-posterior boundaries. These values were then compared with the values obtained via Reader 1 segmentations (Fig. 1).

Anatomical subregions were obtained by first projecting the cartilage onto a two-dimensional plane then dividing it into 12 subregions automatically using previously validated...
techniques (Fig. 4). Segmented cartilage was collapsed onto a single sagittal plane, fit to a cylinder using least squares, and unrolled. In the cylindrical fitting process, the location of each cartilage voxel was expressed in terms of angle, radial distance from the cylinder’s long axis, and slice number. Thirty-five degrees from vertical separated anterior from central subregions. Twenty-five degrees from vertical in the opposite direction separated central from posterior subregions. The anterior-central and central-posterior boundary lines run medial-lateral. The width of the cartilage at the anterior-central boundary line was identified. Any cartilage voxel in the anterior region that was medial to the midpoint of this width was considered to be in the medial anterior subregion and any cartilage voxel in the anterior region that was lateral to this midpoint was considered to be in the lateral anterior subregion. The medial-lateral boundary for the central and posterior regions was obtained by dividing the full width of the segmented cartilage in two equal halves. This boundary separates the two femoral epicondyles. In each slice, cartilage was divided into angular bins of 5°. The superficial and deep halves of each bin in each slice were separated into their respective subregions.

For each subregion, the Spearman correlation and mean absolute error (MAE) between the model’s estimated values and the values derived from Reader 1 were assessed. Reader 2 segmentations were evaluated in the same way, via comparison with Reader 1 segmentations, and these results were compared with the model’s results.

**Comparison of Focal T2 Elevation.** Cluster analysis was performed to identify focal subregions of elevated T2, according to a previously validated approach. T2 maps for the two-dimensional cartilage were calculated for each image, then projected into two dimensions. The projection maps for the two imaging time points of a subject were registered and subtracted to identify the change in T2 value at each pixel location over time. Clusters of contiguous pixels that were all more than one standard deviation above the subject’s mean T2 change across the full cartilage plate were noted. Clusters that covered more than 1% of the area of the cartilage plate were identified and labeled as focal lesions. The cluster intensity and area thresholds remove noise but still identify focal defects. The area of a subject’s femoral cartilage plate covered by clusters was calculated. This has been proposed as a measure of OA risk and progression.

**Statistical Tests.** Agreement between segmentation masks produced by the readers and model was measured using the Dice score and Jaccard index for each test set image. The mean and standard deviation of these metrics were calculated across test set images. Spearman correlation and MAE were used to compare the T2 values and changes in T2 over time derived from segmentations of Reader 1 with those derived from the model. Bland-Altman plots were used to check for bias in the model’s T2 measurements relative to Reader 1 (Supplementary Material 7). P values are reported for each correlation coefficient and each Bland-Altman plot, with 0.05 serving as the threshold for significance.

**Results**

**Training and Evaluation Time**

The final model was trained in 17 epochs, requiring 28 hours. When using the model to segment a new image
volume, it required approximately 6 seconds using an NVIDIA K80 GPU.

**Direct Comparison of Segmentation Masks**

With Reader 1 segmentations used as ground truth, the model had an average volumetric Dice score of 0.85 ± 0.03 and an average volumetric Jaccard index of 0.74 ± 0.04 for the full test set (Table 2). For the test subset segmented by Reader 2, the model had an average volumetric Dice score of 0.85 ± 0.03 and an average volumetric Jaccard index of 0.73 ± 0.05 with regard to Reader 1, while Reader 2 had an average volumetric Dice score of 0.74 ± 0.03 and an average volumetric Jaccard index of 0.59 ± 0.04 with regard to Reader 1. With Reader 2 held as ground truth, the model had an average volumetric Dice score of 0.75 ± 0.03 and an average volumetric Jaccard index of 0.60 ± 0.04.

Volumetric Dice scores and volumetric Jaccard indices were calculated for each test set image. Values reported here are the mean value ± standard deviation calculated across the 28 image volumes in the test set or 14 image volumes in the test subset.

**Comparison of Subregional T2 Mean Values**

The model’s estimates of subregional average T2 values were significantly correlated with those of Reader 1 for all subregions (P < 0.01) except the superficial lateral anterior subregion (P = 0.08). Spearman correlations and the MAE in the model’s estimates of subregional average T2 change are shown in Figure 4 (right). The lateral anterior subregion of one subject was found to have a large, full-thickness lesion that prevented a meaningful analysis of average T2 change for this subregion of this subject and so it was excluded. The other cartilage subregions of this subject were included in the analysis. Eleven of the 12 smallest subregions’ model estimates did not have significant bias (P > 0.05) while the deep medial central subregion had a bias of −2.63 ms (Supplementary Material 7: Bland-Altman Plots).

**Comparison of Longitudinal Subregional T2 Change**

The model’s estimates of four-year change in subregional average T2 values were significantly correlated with those of Reader 1 for all subregions (P < 0.01) except the superficial lateral anterior subregion (P = 0.08). Spearman correlations and the MAE in the model’s estimates of subregional average T2 change are shown in Figure 4 (right). The lateral anterior subregion of one subject was found to have a large, full-thickness lesion that prevented a meaningful analysis of average T2 change for this subregion of this subject and so it was excluded. The other cartilage subregions of this subject were included in the analysis. Eleven of the 12 smallest subregions’ model estimates did not have significant bias (P > 0.05) while the deep medial central subregion had a bias of −2.63 ms (Supplementary Material 7: Bland-Altman Plots).

**Comparison of Focal T2 Elevation**

The percentage of cartilage area affected by significantly increasing T2 clusters estimated using model segmentations correlated with those of Reader 1 with a Spearman correlation of 0.78 (P < 0.01). The MAE in the estimates was 1.7% ± 1.1% of the cartilage plate. The average percentage of cartilage area affected was 8.83% when calculated using Reader 1 segmentations and 8.67% when calculated using model segmentations.

**Discussion**

We have developed a fast, fully automated, end-to-end femoral cartilage segmentation model that agrees with an expert as closely as two experts agree with each other, both in terms of segmentation masks and downstream assessments of cartilage health. The model operates directly on MESE images, eliminating the need to capture a morphological scan in addition to the MESE image to measure T2 values. We have made our model, along with code for replicating the results of this article, publicly available at https://github.com/kathoma/AutomaticKneeMRSegmentation. We have also developed a web application that allows users to drag and drop MESE images into their web browser to have them segmented automatically with no computer programming required. This is available at https://kl.stanford.edu.
Several limitations of our work should be noted. The model’s performance was assessed using semi-automated segmentation refined by an expert as the gold standard, which varies depending on the reader. Also, our test subset that was segmented by Reader 2 was only 14 images. However, the small standard deviation in the Dice scores and mean absolute differences in subregional T2 values observed across those 14 images suggest that these results may be representative of inter-expert agreement more broadly. The model was trained only on subjects without radiographic osteoarthritis, so its performance on images featuring gross morphological disease requires additional investigation. However, a key potential benefit of quantitative MRI is the opportunity to detect OA before it is visible on X-rays. Because analyzed subjects had knee pain and were at risk of developing OA according to OAI criteria, they made an ideal cohort for assessing this use case. Reader 2 did not use the same semi-automated segmentation approach as our gold standard, but instead manually segmented the images. Both semi-automated and manual...
segmentation are commonly used for knee cartilage in MESE images, so comparisons between these approaches provide valuable context for assessing the reproducibility of our model relative to different standard practices. The fact that our model agrees with Reader 1 segmentations more closely than Reader 2 segmentations suggests that our model has learned to replicate the nuanced, fine details of Reader 1 that other readers may disagree with. However, this is an issue that currently limits comparisons between any two studies in the body of knee cartilage literature that use different readers to segment their images. By making our model publicly available, we introduce the potential to enhance comparisons between future studies that leverage our model in their work.

Other works have aimed to automatically segment T2-weighted knee MRIs using atlas-based approaches,

![Figure 4. Model’s agreement with gold standard in its estimates of subregional mean T2 values (left) and in its estimates of four-year change in subregional mean T2 values (right). Absolute errors are reported as mean ± standard deviation. All correlations are significant ($P < 0.01$) except for the four-year change in the superficial lateral anterior subregion ($P = 0.08$, denoted with *). Estimates of subregional T2 values all have a mean absolute error (MAE) less than 2 ms and estimates of four-year change all have an MAE less than 3 ms. This is comparable to inter-reader agreement.](image-url)
shape models, and deep learning. One work developed a deep learning model for fat-suppressed T2-weighted fast spin-echo images. They report a femoral cartilage segmentation Dice score of 0.81 ± 0.04, similar to our model's 0.85 ± 0.03. Although they used T2-weighted images, these images do not provide quantitative T2 maps like the MESE images we used.

In a prior work that used all MESE images from the OAI baseline cohort, images were segmented via multiple steps of non-rigid registrations with an atlas image. Relative to our approach, this is significantly more time-consuming. In estimates of the T2 value in anatomical regions of the femoral cartilage, they reported a MAE of 2.16 ms for the lateral femur (Pearson correlation $R = 0.82$) and 1.73 ms for the medial femur (Pearson correlation $R = 0.75$). In contrast, our model has a MAE of 0.61 ms for the lateral femur (Pearson correlation $R = 0.96$) and 0.61 ms for the medial femur (Pearson correlation $R = 0.96$) (Supplementary Material 5: Pearson Correlation Coefficients). They report an average bias of −1.2 while our bias for the medial femur is −0.1 and for the lateral femur is 0.06 (Supplementary Material 7: Bland-Altman Plots). They did not report the demographics or OA severities of the subjects in their test set, which may affect these comparisons. However, no subjects in our test set had a four-year change in lateral femoral cartilage T2 less than our model's MAE of 0.61 ms, but 12 of 14 subjects had a 4-year change larger than our model's MAE 2.16 ms. Similarly, 1 of 14 subjects had a four-year change in medial femoral cartilage T2 less than our model's MAE of 0.61 ms, but 12 of 14 subjects had a four-year change less than their model's MAE 1.73 ms. This suggests that our model provides clinically significant improvements in the ability to detect longitudinal change.

A third work combined a deep learning model with a 3-dimensional simplex deformable model to segment MESE images. Their simplex deformable modeling step required 56% of their algorithm’s total segmentation computation time, which our method avoids. Their approach used T2 maps as input to their segmentation system, in contrast to our use of the second echo image. By eliminating the need to calculate T2 values for all noncartilage voxels, our approach also reduces the computational time of this step by approximately two orders of magnitude. These time-saving features come with no cost in model performance. From their reported results, it can be derived that their model had a MAE of 1.4 ms in estimating the average T2 value of the full femoral cartilage plate. This is higher than our MAE of 0.41 ± 0.32 ms. It can also be derived that their Jaccard index was 0.75 ± 0.06, comparable to our model’s 0.74 ± 0.04.

It is important to note that the images and researchers performing reference segmentations differed between our work and each of these prior works, for which models are not publicly available. This precludes ideal comparisons of accuracy and performance between models. We have therefore made our model and segmentation data publicly available to facilitate future comparisons.

Our model’s performance may be near the limit of reproducibility for measuring regional T2 values. In an assessment of scan-rescan reproducibility of T2 values in healthy controls imaged twice on the same day, it was found that the medial and lateral femoral cartilage estimates’ root mean squared coefficient of variation (RMS-CV) were between 4.0% and 4.5%. In contrast, our model’s RMS-CV in estimating T2 values were 1.46% for the lateral femoral cartilage and 1.48% for the medial femoral cartilage (Supplementary Material 6). In another study, inter-reader RMS-CV for whole knee cartilage mean T2 was 1.57%, while our model’s is 1.0%. Our model reproduces the measurements of an expert on individual images more closely than a single experts’ measurements of healthy controls agree with one another in the scan-rescan context and with a similar level of agreement as two experts reading the same image.

Although Dice scores and average T2 values for large cartilage regions are useful for assessing model performance, we go further. First, we assessed model performance using smaller subregions of the femoral cartilage. It is more difficult for two readers (or a reader and model) to have strong agreement on the average T2 value in smaller subregions because the mean is calculated over fewer voxels. Yet it is important to assess how well models agree with experts on the small subregions that result from splitting the cartilage along medial-lateral, superficial-deep, and anterior-central-posterior boundaries because this is frequently how T2 values are tracked in research. Second, we assessed how well the model is able to capture changes in T2 value over time for each subregion. T2 values are known to vary significantly across individuals, including healthy controls. The efficacy of treatments is therefore often assessed by the change in patients’ T2 value over time. We measure how well our model tracks longitudinal T2 change in two ways: (1) calculating changes in subregional mean T2 values over time and (2) tracking the extent of focal areas of T2 worsening. While agreement between expert readers regarding these longitudinal metrics is not assessed here and is scarce in prior research, the error in the model’s estimates of subregional T2 change over four years are similar to the disagreement in subregional T2 value between our two readers for single time points.

In conclusion, we present an open source, fast and fully automated femoral cartilage segmentation model that agrees with experts as closely as experts agree with one another. This makes it possible to leverage MESE-based findings in large-scale studies and has the potential to unlock new lines of inquiry on the earliest stage of OA.
Acknowledgments and Funding
This work was supported by the National Institutes of Health (NIH) Big Data to Knowledge (BD2K) Research Grant U54EB020405 and NIH Grant P41 EB027060. The funding source played no role in the research.

Declaration of Conflicting Interests
The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Dr. Chaudhari reports personal fees from Skope MR, personal fees and equity from Subtle Medical, personal fees from Chondrometrics GmbH, personal fees from Image Analysis Group, personal fees from Edge Analytics, personal fees from Culvert Engineering, personal fees from ICM Co., equity from Brain Key, equity from LVIS Corp, and grants from GE Healthcare. All of these activities are outside the scope of the submitted work.

Ethical Approval
Our analysis was approved by our institution’s institutional review board (eProtocol #31592).

Informed Consent
All subjects provided written informed consent.

Trial Registration
Not applicable.

ORCID iD
Kevin A. Thomas https://orcid.org/0000-0002-3546-330X

References
1. Jordan JM, Helmick CG, Renner JB, Luta G, Dragomir AD, Woodard J, et al. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol. 2007;34(1):172-80.
2. Felson DT, Naimark A, Anderson J, Kazis L, Castelli WP, Meenan RF. The prevalence of knee osteoarthritis in the elderly. The Framingham Osteoarthritis Study. Arthritis Rheum. 1987;30(8):914-8.
3. Bagge E, Bjelle A, Valkenburg HA, Svanborg A. Prevalence of radiographic osteoarthritis in two elderly European populations. Rheumatol Int. 1992;12(1):33-8.
4. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthritis. Ann Rheum Dis. 1957;16(4):494-502. doi:10.1136/ard.16.4.494
5. Guermazi A, Roemer FW, Burstein D, Hayashi D. Why radiography should no longer be considered a surrogate outcome measure for longitudinal assessment of cartilage in knee osteoarthritis. Arthritis Res Ther. 2011;13(6):247.
6. Andriacchi TP, Mundermann A, Smith RL, Alexander EJ, Dyrbj CO, Koo S. A framework for the in vivo pathomechanics of osteoarthritis at the knee. Ann Biomed Eng. 2004;32(3):447-57.
7. Eckstein F, Burstein D, Link TM. Quantitative MRI of cartilage and bone: degenerative changes in osteoarthritis. NMR Biomed. 2006;19(7):822-54.
8. Mosher TJ, Dardzinski BJ. Cartilage MRI T2 relaxation time mapping: overview and applications. Semin Musculoskelet Radiol. 2004;8(4):355-68.
9. Dunn TC, Lu Y, Jin H, Ries MD, Majumdar S. T2 relaxation time of cartilage at MR imaging: comparison with severity of knee osteoarthritis. Radiology. 2004;232(2):592-8.
10. Chaudhari AS, Kogan F, Pedaio V, Majumdar S, Gold GE, Hargreaves BA. Rapid knee MRI Acquisition and analysis techniques for imaging osteoarthritis. J Magn Res Imaging. 2020;52(5):1321-39. doi:10.1002/jmri.26991
11. Stehling C, Baum T, Mueller-Hoecker C, Liebl H, Carballido-Gamio J, Joseph GB, et al. A novel fast knee cartilage segmentation technique for T2 measurements at MR imaging—data from the Osteoarthritis Initiative. Osteoarthritis Cartilage. 2011;19(8):984-9.
12. Jordan CD, McWalter EJ, Monu UD, Watkins RD, Chen W, Bangerter NK, et al. Variability of CubeQuant T1ρ, quantitative DESS T2, and cones sodium MRI in knee cartilage. Osteoarthritis Cartilage. 2014;22(10):1559-67.
13. Duryea J, Iranpour-Boroujeni T, Collins JE, Vanwymgaarden C, Guermazi A, Katz JN, et al. Local area cartilage segmentation: a semiautomated novel method of measuring cartilage loss in knee osteoarthritis. Arthritis Care Res (Hoboken). 2014;66(10):1560-5.
14. Schaefer LF, Sury M, Yin M, Jamieson S, Donnell I, Smith SE, et al. Quantitative measurement of medial femoral knee cartilage volume—analysis of the OA Biomarkers Consortium FNIH Study cohort. Osteoarthritis Cartilage. 2017;25(7):1107-13.
15. Williams AA, Titchenal MR, Do BH, Guha A, Chu CR. MRI UTE-T2* shows high incidence of cartilage subsurface matrix changes 2 years after ACL reconstruction. J Orthop Res. 2019;37(2):370-7.
16. Chaudhari AS, Black MS, Eijgenraam S, Wirth M, Maschek S, Sveinsson B, et al. Five-minute knee MRI for simultaneous morphometry and T2 relaxometry of cartilage and meniscus and for semiquantitative radiological assessment using double-echo in steady-state at 3T. J Magn Res Imaging. 2018;47(5):1328-41.
17. Eijgenraam SM, Chaudhari AS, Reijman M, Bierma-Zeinstra SMA, Hargreaves BA, Runhaar J, et al. Time-saving opportunities in knee osteoarthritis: T2 mapping and structural imaging of the knee using a single 5-min MRI scan. Eur Radiol. 2020;30(4):2231-40.
18. Heimann T, Morrison BJ, Styner MA, Niethammer M, Warfield SK, editors. Proceedings of the MICCAI Workshop on Medical Image Analysis for the Clinic. CreateSpace; 2010:207-14.
19. Norman B, Pedoia V, Majumdar S. Use of 2D U-Net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry. Radiology. 2018;288(1):177-85. doi:10.1148/radiol.2018172322
20. Tamez-Peña JG, Farber J, González PC, Schreyer E, Schneider E, Totterman S. Unsupervised segmentation and quantification of anatomical knee features: data from the Osteoarthritis Initiative. IEEE Trans Biomed Eng. 2012;59(4):1177-86.

21. Dam EB, Lillholm M, Marques J, Nielsen M. Automatic segmentation of high- and low-field knee MRIs using knee image quantification with data from the osteoarthritis initiative. J Med Imaging (Bellingham). 2015;2(2):024001.

22. Raj A, Vishwanathan S, Ajani B, Krishnan K, Agarwal H. Automatic knee cartilage segmentation using fully volumetric convolutional neural networks for evaluation of osteoarthritis. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); April 4-7, 2018; Washington, DC.

23. Wirth W, Eckstein F, Kemnitz J, Baumgartner CF, Konukoglu E, Fuerst D, et al. Accuracy and longitudinal reproducibility of quantitative femorotibial cartilage measures derived from automated U-Net-based segmentation of two different MRI contrasts: data from the osteoarthritis initiative healthy reference cohort. MAGMA. 2021;34(3):337-54. doi:10.1007/s10334-020-00889-7

24. Fürst D, Wirth W, Chaudhari A, Eckstein F. Layer-specific analysis of femorotibial cartilage T2 relaxation time based on registration of segmented double echo steady state (DESS) to multi-echo-spin-echo (MESE) images. MAGMA. 2020;33(6):819-28. doi:10.1007/s10334-020-00852-6

25. Urish KL, Keffalas MG, Durkin JR, Miller DJ, Chu CR, Mosher TJ. T2 texture index of cartilage can predict early symptomatic OA progression: data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2013;21(10):1550-7.

26. Zhong H, Miller DJ, Urish KL. T2 map signal variation predicts symptomatic osteoarthritis progression: data from the Osteoarthritis Initiative. Skeletal Radiol. 2016;45(7):909-13.

27. Fripp J, Crozier S, Warfield SK, Ourselin S. Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. IEEE Trans Med Imaging. 2010;29(1):55-64.

28. Juras V, Szomolanyi P, Schreiner MM, Unterberger K, Kurekova A, Hager B, et al. Reproducibility of an automated quantitative MRI assessment of low-grade knee articular cartilage lesions. Cartilage. Published online September 29, 2020. doi:10.1177/1947605320961165

29. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis Cartilage. 2008;16(12):1433-41.

30. Pdeoia V, Lee J, Norman B, Link TM, Majumdar S. Diagnosing osteoarthritis from T2 maps using deep learning: an analysis of the entire Osteoarthritis Initiative baseline cohort. Osteoarthritis Cartilage. 2019;27(7):1002-10.

31. Liu F, Zhou Z, Jang H, Samsonov A, Zhao G, Kijowski R. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. Magn Reson Med. 2018;79(4):2379-91.

32. Liu F, Chaudhary R, Hurley SA, Del Rio AM, Alexander AL, Samsonov A, et al. Rapid multicomponent T2 analysis of the articular cartilage of the human knee joint at 3.0T. J Magn Reson Imaging. 2013;39(5):1124-32.

33. Halilaj E, Hastie TJ, Gold GE, Delp SL. Physical activity is associated with changes in knee cartilage microstructure. Osteoarthritis Cartilage. 2018;26(6):770-4.

34. Wirth W, Masechek S, Roemer FW, Sharma L, Duda GN, Eckstein F. Radiographically normal knees with contralateral joint space narrowing display greater change in cartilage transverse relaxation time than those with normal contralateral knees: a model of early OA? Data from the Osteoarthritis Initiative (OAI). Osteoarthritis Cartilage. 2019;27(11):1663-8.

35. Joseph GB, McCulloch CE, Nevitt MC, Neumann J, Gersing AS, Kretzschmar M, et al. Tool for osteoarthritis risk prediction (TOARP) over 8 years using baseline clinical data, X-ray, and MRI: Data from the osteoarthritis initiative. J Magn Reson Imaging. 2018;47(6):1517-26.

36. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. arXiv [csCV]. Available from: http://arxiv.org/abs/1505.04597

37. Smith HE, Mosher TJ, Dardzinski BJ, Collins BG, Collins CM, Yang QX, et al. Spatial variation in cartilage T2 of the knee. J Magn Reson Imaging. 2001;14(1):50-5.

38. Monu UD, Jordan CD, Samuelson BL, Hargreaves BA, Gold GE, McWalter EJ. Cluster analysis of quantitative MRI T2 and T1ρ relaxation times of cartilage identifies differences between healthy and ACL-injured individuals at 3T. Osteoarthritis Cartilage. 2017;25(4):513-20.

39. Crowder IA, Mazzoli V, Black MS, Watkins LE, Kogan F, Hargreaves BA, et al. Characterizing the transient response of knee cartilage to running: Increases in cartilage T2 of female recreational runners. J Orthop Res. Published online January 22, 2021. doi:10.1002/jor.24994

40. Liu F, Zhou Z, Samsonov A, Blankenbaker D, Larison W, Kanarek A, et al. Deep learning approach for evaluating knee MR images: achieving high diagnostic performance for cartilage lesion detection. Radiology. 2018;298(1):160-9.

41. Li X, Pdeoia V, Kumar D, Rivoire J, Wyatt C, Lansdown D, et al. Cartilage T1ρ and T2 relaxation times: longitudinal reproducibility and variations using different coils, MR systems and sites. Osteoarthritis Cartilage. 2015;23(12):2214-23.

42. Bittersohl B, Hosalkar HS, Sondern M, Miese FR, Antoch G, Krauspe R, et al. Spectrum of T2* values in knee joint cartilage at 3 T: a cross-sectional analysis in asymptomatic young adult volunteers. Skeletal Radiol. 2014;43(4):443-52.

43. Liu F, Chaudhary R, Hurley SA, Del Rio AM, Alexander AL, Samsonov A, et al. Rapid multicomponent T2 analysis of the articular cartilage of the human knee joint at 3.0T. J Magn Reson Imaging. 2013;39(5):1191-7.

44. Souza RB, Kumar D, Calixto N, Singh J, Schooler J, Subburaj K, et al. Response of knee cartilage T1rho and T2 relaxation times to in vivo mechanical loading in individuals with and without knee osteoarthritis. Osteoarthritis Cartilage. 2014;22(10):1367-76.