

Osteoarthritis and Cartilage



T2 texture index of cartilage can predict early symptomatic OA progression: data from the osteoarthritis initiative

K.L. Urish †*, M.G. Keffalas ‡, J.R. Durkin §, D.J. Miller ||, C.R. Chu ¶, T.J. Mosher #

† Department of Orthopaedics and Rehabilitation, Division of Musculoskeletal Sciences, College of Medicine, The Pennsylvania State University, 30 Hope Drive, EC089, Hershey, PA 17033, USA

‡ Department of Electrical Engineering, The Pennsylvania State University, 227C Electrical Engineering West, University Park, PA 16802, USA

§ School of Medicine, University of Pittsburgh, 533 Scaife Hall, Pittsburgh, PA 15260, USA

|| Department of Electrical Engineering, The Pennsylvania State University, 227C Electrical Engineering West, University Park, PA 16802, USA

¶ Department of Orthopaedic Surgery, Stanford University Medical Center, 450 Broadway St, MC6342, Redwood City, CA 94063, USA

Department of Radiology, Penn State Milton S. Hershey Medical Center, 500 University Drive, MC, H066, Hershey, PA 17033, USA

ARTICLE INFO

Article history:

Received 30 December 2012

Accepted 5 June 2013

Keywords:

Cartilage T2

T2 heterogeneity

Osteoarthritis

MRI

Image biomarkers

Signal texture

SUMMARY

Objective: There is an interest in using Magnetic Resonance Imaging (MRI) to identify pre-radiographic changes in osteoarthritis (OA) and features that indicate risk for disease progression. The purpose of this study is to identify image features derived from MRI T2 maps that can accurately predict onset of OA symptoms in subjects at risk for incident knee OA.

Methods: Patients were selected from the Osteoarthritis Initiative (OAI) control cohort and incidence cohort and stratified based on the change in total Western Ontario and McMaster Universities Arthritis (WOMAC) score from baseline to 3-year follow-up (80 non-OA progression and 88 symptomatic OA progression patients). For each patient, a series of image texture features were measured from the baseline cartilage T2 map. A linear discriminant function and feature reduction method was then trained to quantify a texture metric, the T2 texture index of cartilage (TIC), based on 22 image features, to identify a composite marker of T2 heterogeneity.

Results: Statistically significant differences were seen in the baseline T2 TIC between the non-progression and symptomatic OA progression populations. The baseline T2 TIC differentiates subjects that develop worsening of their WOMAC score OA with an accuracy between 71% and 76%. The T2 TIC differences were predominantly localized to a dominant knee compartment that correlated with the mechanical axis of the knee.

Conclusion: Baseline heterogeneity in cartilage T2 as measured with the T2 TIC index is able to differentiate and predict individuals that will develop worsening of their WOMAC score at 3-year follow-up.

© 2013 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Introduction

Plain radiography is the imaging standard used to diagnose osteoarthritis (OA). Progression can be tracked following serial radiographs by assessing changes in joint space width and the appearance of other OA hallmarks, including osteophytes¹.

However, radiographs are an unsatisfactory imaging modality for OA because it lacks the sensitivity to capture and monitor early disease progression when intervention has the greatest potential for modifying patient outcomes. Ultimately, there is a poor correlation between radiographic OA changes, clinical complaints of pain², and disease progression. Because Magnetic Resonance Imaging (MRI) has the capability to directly image cartilage, it has the potential to provide sensitive and specific measurements of tissue damage occurring at an early stage of OA. There have been preliminary applications of compositional MRI techniques to detect changes in water and proteoglycan content and anisotropy of collagen fibers^{3–6} associated with early degradation⁷. New imaging modalities are needed that can detect early changes in OA.

The MRI transverse relaxation time (T2) of articular cartilage is a quantitative tissue parameter that is strongly dependent on the

* Address correspondence and reprint requests to: K.L. Urish, Department of Orthopaedics and Rehabilitation, Division of Musculoskeletal Sciences, College of Medicine, The Pennsylvania State University, 30 Hope Drive EC089, Hershey, PA 17033, USA. Tel: 1-412-736-4261; Fax: 1-717-531-7583.

E-mail addresses: kurish@hmc.psu.edu, ken.urish@gmail.com (K.L. Urish), mkeffalas@yahoo.com (M.G. Keffalas), john.r.durkin@gmail.com (J.R. Durkin), djmiller@enr.psu.edu (D.J. Miller), chucr@stanford.edu (C.R. Chu), tmosher@hmc.psu.edu (T.J. Mosher).

structure of the extracellular type II collagen matrix^{4,8} and cartilage water content^{4,7,9}. In healthy tissue the T2 distribution of knee articular cartilage has a well-recognized spatial pattern in T2 values that reflect regional variation in the collagen fiber anisotropy and water content. This pattern is observed in T2-weighted clinical MR images, where low signal is observed near bone, gradually increasing in signal intensity toward the articular surface reflecting the depth dependency in orientation of the collagen matrix with respect to the applied magnetic field. Loss of this pattern is observed and used to identify focal cartilage injury. Early in the pathogenesis of OA, loss of cartilage anisotropy and increase in water content produce greater variability in T2 values between neighboring voxels. With more chronic cartilage damage, areas of low signal intensity may be observed near sites of focal cartilage injury, producing greater variation in the pattern of T2 signal in cartilage¹⁰.

We postulate that disruption of the normal spatial variation in cartilage T2 is an early indication of cartilage injury that may predict individuals at risk for developing OA symptoms. To test this hypothesis, we analyzed data obtained from the Osteoarthritis Initiative (OAI) data set to identify quantitative measures of cartilage T2 heterogeneity that are present at baseline in subjects that subsequently develop OA symptoms.

Materials and methods

Population cohort

Data was obtained from the OAI database, available for public access at <http://www.oai.ucsf.edu/>. Specific datasets used from the OAI were kXR SQ reading (BU) [version 0.5] for measurement of the Kellgren and Lawrence (KL) grade scores. Clinical symptoms were assessed with the Western Ontario and McMaster Universities Arthritis (WOMAC) questionnaire at the time of magnetic resonance screening¹¹.

A total of 168 patients were selected from the OAI prospective cohort ($n = 4,796$); 80 were selected from the unexposed control subcohort for the non-progression population and 88 from the incidence subcohort for the symptomatic progression population (Fig. 1). Non-progression subjects were selected from the OAI unexposed control subcohort, defined at baseline and at 3 years by a WOMAC score <10 with a KL score ≤ 2 (KL = 0, $n = 72$; KL = 1, $n = 8$), and no risk factors for OA progression; The symptomatic OA progression cohort was selected from the OAI incidence subcohort based on the initial (baseline) criteria of a WOMAC score ≤ 10 , but with a change in WOMAC score of >10 within 3 years from baseline, and minimal baseline radiographic signs of OA defined as a KL ≤ 2 (KL = 0, $n = 41$; KL = 1, $n = 25$; KL = 2, $n = 22$). Analysis of this group was repeated with a more restrictive exclusion criteria of no radiographic signs of OA defined as a KL <2 (KL = 0, $n = 41$; KL = 1, $n = 25$). Exclusion criteria for the entire OAI cohort included rheumatoid arthritis, bilateral total knee joint replacement, and a positive pregnancy test. Institutional review board approval was obtained at all participating institutions in the OAI, and informed consent was obtained by all participants in the study¹².

MR image acquisition, plain radiographic, and clinical assessment

In the OAI cohort, 3 dimensional sagittal dual echo steady state (DESS) and T2 mapping images were acquired from the imaging database freely available by request (<http://oai.epi-ucsf.org/>)¹³. Standard bilateral standing posterior–anterior fixed flexion knee radiographs were obtained at the baseline visit. Knee radiographs were graded using the KL scoring system¹⁴. The mechanical axis was determined using the standard technique of measuring the

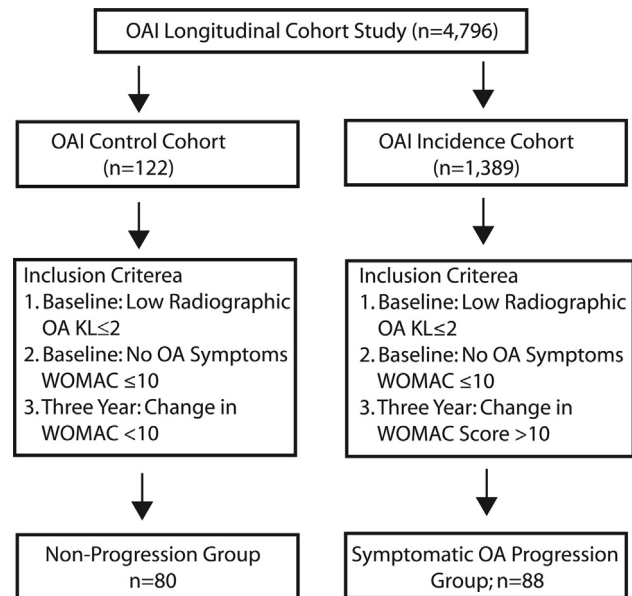


Fig. 1. Experiment design schematic. The non-progression group was collected from the OAI control cohort ($n = 80$). The rapid progression population was collected from the OAI incidence cohort ($n = 88$). At the initial time point, the progression population was asymptomatic, and at the 3-year time point, this population experienced a WOMAC change >10 . Segmentation and registration software were used to extract texture features from baseline T2 maps. The group was divided into separate, independent training and testing subsets. An image classifier, SVM, was used to develop a texture metric (T2 TIC) to predict OA progression using the training subset. MFE was then used for feature reduction. The accuracy of T2 TIC was measured on the independent testing subset.

angle placed from the center of the femoral head to the medial tibial prominence to the midline of the ankle¹⁵. OAI datasets used included the baseline and 1-year imaging data set 0.E.1 and 0.C.2.

Image registration, segmentation, and T2 maps

DESS and T2 images were registered using the Mattes mutual information metric¹⁶. Validation, accuracy, and precision of the registration process have been previously described¹⁷. Registration software was built using the insight toolkit, a C++ open source image analysis library (www.itk.org)¹⁸. The software is freely available (www.imageK.org).

Segmentation was completed on DESS images. Segmentation of the femoral and patellar cartilage was completed using custom semi-automated software implementing a global active statistical shape model with a local active contour model as previously described¹⁹. Gross inaccuracies in the segmentation were manually corrected. The segmentation was completed by a single individual, and was completed in approximately 20 min per patient.

Binary masks of the lateral and medial femoral condyle and patella were generated from the segmented images. There were 11 regions of interest (ROI) identified per individual. The lateral and medial masks were split into five sections for each individual. The division between the medial and lateral compartment was defined as the midpoint in sequences, which roughly correlated with the anterior cruciate ligament, and was completed manually. These five subdivisions in the middle and lateral compartments are based on equal sagittal divisions defined in the region between the anterior and posterior apex of the femoral cartilage on the femoral condyles (Fig. S1). The patella region was treated as a single section.

T2 maps were calculated from the Multi-Slice-Multi-Echo T2 images. Calculation of the T2 maps has been previously described²⁰.

Briefly, the T2 maps are calculated on a voxel-by-voxel basis using a linear least squares fitting. The MR T2 signal decay of cartilage is mono-exponential, and the signal intensity decay can be expressed as an exponential decay as a function of time for each voxel. Masks of the T2 maps were created from the DESS segmentations after registration.

Image feature extraction

We chose as candidate features some well-known descriptors of image texture (local entropy, variance, cross-correlation, run-lengths, histogram-based features); and integrated a feature reduction step within the classifier training. Candidate features were calculated from each T2 map using the segmented binary masks ROI using a Matlab script (Mathworks, Natick, MA). Each feature was independently measured in each of the 11 sections on each knee. There were four main categories of features: histogram, gray level co-occurrence matrix (GLCM), gray level run length matrix (GLRL), and z-score. The numbers reported below are the totals from all 11 sections. A 32-bin histogram was used to calculate the mean, variance, entropy, and central moments^{21,22}. GLCM features were calculated from the GLCMs at unit distance and angles 0, 45, 90, 135°, and 90° in the z direction²³. GLRL features were calculated from GLRLs at angles 0 and 90°²⁴. The z-score was calculated for all voxels in each section²⁵. The mean value, variance, minimum value, maximum value, and range of values were then calculated ($n = 55$). Over all of the 11 sections, a total of 725 features were measured on each T2 map. All features were normalized to the range $[-1,1]$.

Statistics

Classification

An image classifier, support vector machine (SVM), was used to quantify differences in the texture metric and develop a model to predict OA progression by classifier training and testing based on dividing the sub-populations into training and test sets (Fig. S2). SVM training and testing were implemented using the LIBSVM Matlab interface²⁶. To assess the performance of the classifier, we randomly divided the entire cohort into equal-sized training (non-progression, $n = 40$; symptomatic OA progression, $n = 44$) and test (non-progression, $n = 40$; symptomatic OA progression, $n = 44$) subsets with equal numbers of non-progression and symptomatic OA progression individuals. This was repeated to create one thousand sets of corresponding independent training and testing subsets where no patients from the training set were included in the matching testing set. In each of the 1,000 trials, the SVM classifier was trained to discriminate between non-progression and symptomatic OA progression populations using all 725 features on the training set, and the accuracy of the classifier and confusion matrix was measured on the independent test set. It should be emphasized that the training and test subsets were independent, and measurements of the accuracy of the model did not include any patients from the training set used to build the model.

Feature elimination

Margin-based feature elimination (MFE) was used to eliminate redundant and uninformative candidate features^{27,40–44} (Supplemental Methods). For each trial, SVM training was coupled with MFE to identify a reduced set of essential features. The accuracy of the reduced feature set was tested on the test data set, and the confusion matrix was again determined. After classification was completed and the T2 TIC was calculated, the T2 TIC of each compartment was determined. Results were normalized based on the number of regions in each compartment, and averaged across the separate trials.

Partial sum measurements

The contribution of the medial and lateral compartment to the overall T2 TIC was determined using a partial weighted sum technique. After classification was completed and the T2 TIC was calculated (the SVM score), the T2 TIC of each separate compartment was determined. Only the symptomatic progression population was included in the analysis as we were investigating the mechanical axis alignment's contribution to symptomatic OA progression. In each of the trials, the partial weighted linear sum of the medial femoral condyle, lateral femoral condyle, and patella contribution to each individual's overall SVM score was determined. Results were normalized based on the number of regions in each compartment (5 for the medial and lateral condyle, one for the patella), averaged across the trials, and rank ordered.

Data is expressed as mean \pm 95% confidence interval, except where noted. Direct comparisons between two populations were made using a two-tailed Student *t*-test. Statistical significance was determined if $P < 0.05$. Multiple group comparisons were made using two-way analysis of variance, using the Student–Newman–Keuls pairwise comparison to determine significance levels. Conventions for box plot include the middle band representing the median of the population, where the diagonal lines (notches) represent an approximate 95% confidence interval of the median. A lack of overlap between the notches of the two populations suggests that the two medians are statistically different. The bottom and top of the boxes represent the 25% and 75% quantiles of the data. The whiskers represent the minimum and maximum value of the population. Outliers are shown as circular glyphs. Receiver operating characteristic (ROC) analysis was performed on the entire set using standard techniques²⁸.

Results

Demographics

We defined two populations in the OAI cohort: a non-progression and symptomatic OA progression population. These populations were comparable with regard to age, sex, and Body Mass Index (BMI). As expected by cohort definitions, the non-progression population had lower WOMAC and KL scores as compared to the symptomatic OA progression population; however, the values are clinically comparable (Table I). The yearly change in WOMAC scores in the symptomatic OA progression population was highly variable while there was little change in the non-progression group (Table I). The incidence of any reported traumatic event to the knee was 17% in the symptomatic OA progression group and 4% in the non-progression group. The group of

Table I

Demographic data of the control and symptomatic OA progression cohorts. The change in WOMAC score is over a 3-year period.

	Progression group	Control group	<i>P</i> value
No. of patients	88	80	
Mean age (year)	56.0 [54.2, 57.8]	54.3 [52.8, 55.8]	0.14
M:F ratio	0.48	0.74	0.15
Mean BMI (kg/m ²)	25.3 [24.4, 26.3]	24.7 [24.0, 225.4]	0.32
Mean KL	0.8 [0.6, 1.0]	0.1 [0, 0.2]	<0.01
Mean KL 3-year Δ	0.4 [0.2, 0.5]	0.1 [0, 0.2]	<0.01
Mean initial WOMAC	3.5 [2.8, 4.1]	0.2 [0.1, 0.4]	<0.01
Mean minimum WOMAC Δ /year	-4.6 [-6.3, -3.0]	-0.8 [-1.2, -0.4]	<0.01
Mean maximum WOMAC Δ /year	18.2 [16.4, 20.0]	0.7 [0.3, 1.1]	<0.01
Mean WOMAC Δ /year	6.5 [5.9, 7.2]	0.0 [-0.1, 0.0]	<0.01
Mean WOMAC 3-year Δ	19.6 [17.8, 21.5]	-0.1 [-0.2, 0.1]	<0.01

Range represents 95% confidence interval; Δ = change.

patients that reported a traumatic event in the symptomatic OA progression population continued to have a large variation in yearly changes in their WOMAC score at the other measured time points.

T2 TIC prediction of symptom development

The T2 TIC had good separation of the symptomatic OA progression and non-progression populations. An image classifier, SVM, was used to quantify differences in the texture metric (texture index of cartilage (TIC)) and develop a model to predict OA progression by classifier training and testing based on dividing the sub-populations into training and test sets. It should be emphasized that the training and test subsets were independent, and measurements of the accuracy of the model did not include any images from the training set used to build the model. By definition, the classifier sets a signal texture index value of zero as the decision boundary so that any positive value is classified as OA progression and any negative value is classified as a control. Comparison of the histogram and box plots of these two populations demonstrates these differences are statistically significant [Fig. 2(A and B)].

The T2 TIC can be used as a prognostic image biomarker for worsening WOMAC score. Three separate cases of classifier accuracy were analyzed. First, the accuracy based on using the entire set of all 725 features, before feature elimination, was measured. The balanced accuracy of the classifier was $76.2 \pm 0.7\%$, corresponding to an average sensitivity of $74.1 \pm 0.4\%$ and an average specificity of $78.4 \pm 0.6\%$. Second, MFE was used to remove redundant and uninformative features, significantly reducing the feature space. An average of only 20 of the 725 features was needed to maintain a comparable level of accuracy. The balanced accuracy of the system with was $71.7 \pm 0.3\%$, corresponding to an average sensitivity of $73.3 \pm 0.5\%$ and an average specificity of $70.2 \pm 0.5\%$. ROC analysis showed excellent classifier performance, and tradeoffs between specificity and sensitivity as a function of the SVM decision boundary (Fig. 3). A problem with this approach is that a new unique features set is generated on each trial (Fig. S2). At a sacrifice of some minimal bias to obtain a single feature set, MFE can be conducted simultaneously across all testing sets. When texture features are eliminated simultaneously on each of the test trials to generate a single set of features, balanced accuracy was $71.2\% \pm 0.3\%$ with a sensitivity of $72.3\% \pm 0.5\%$ and a specificity of $70.1 \pm 0.5\%$. Only 22 texture features were necessary to build the T2 TIC from the 725 initial MRI signal texture features measured using this method. Exclusion of patients with a radiographic KL grade of 2

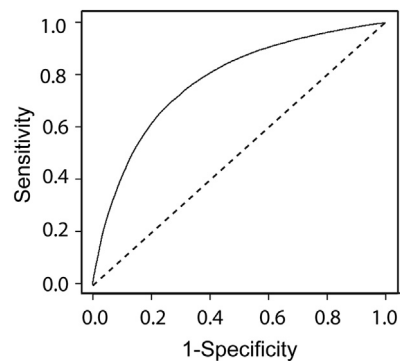


Fig. 3. ROC analysis shows the prognostic accuracy of the T2 TIC has reasonable accuracy. The sensitivity is the true positive rate. The specificity is equivalent to one minus the false positive rate. The diagonal line indicates the result of random chance.

from the symptomatic OA progression cohort resulted in a minor loss of accuracy. After MFE was used to remove redundant features, balanced accuracy was measured at $69.3 \pm 0.3\%$.

The remainder of the discussion will focus on the analysis of the single feature set of the original cohort across the entire testing set. A variety of different features were used in the T2 TIC metric and were associated with specific cartilage sections (Table SI). The lateral compartment contributed 69% of the features and the medial compartment contributed 27% with the remainder of the features from the patella.

T2 TIC is associated with a dominant knee compartment

The image texture features that predict symptomatic progression of OA for most individuals are primarily localized within a single compartment of the knee. This is demonstrated as follows. The T2 TIC is calculated from a weighted sum of image feature measurements from the lateral and medial compartments and the patella. By separately considering the features from each compartment (lateral, medial, patella) and finding the partial sum of the SVM score for each section from the overall SVM score for the knee, the effective contribution from each compartment to the overall decision can be determined. The symptomatic OA progression population was considered separately in this analysis. The contribution of features in each compartment to the overall T2 TIC shows substantial separation between each compartment

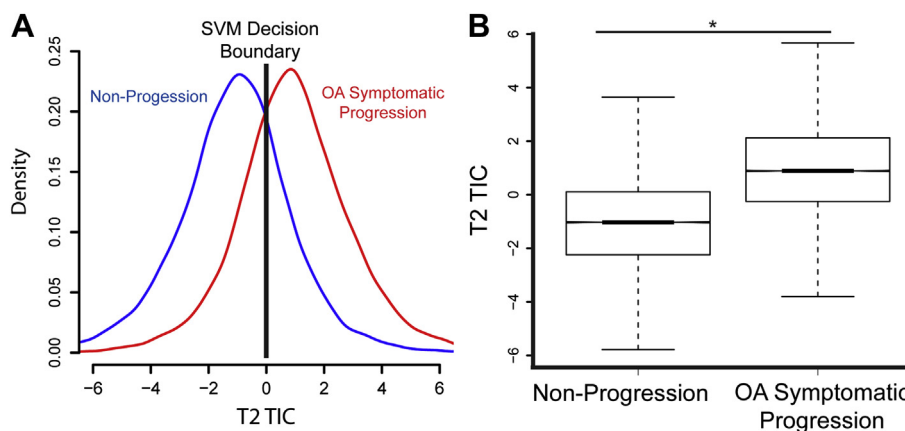


Fig. 2. T2 TIC identifies early signs of OA on T2 maps. (A) Histograms of the T2 TIC for the non-progression and OA progression populations. A positive score corresponds to a decision of symptomatic OA, and a negative score indicates a decision of asymptomatic. Accuracy is 76%. The SVM decision boundary is indicated by the black line. Results shown are based on 1,000 trials, with on average 22 features needed to build the T2 TIC. (B) Notched box plots comparing the control and OA populations. The notched line is visualized as the bold line at the median as the confidence interval is small. $*P < 0.05$.

[Fig. 4(A)] and the means of these compartments are statistically different [Fig. 4(B)].

To test the observation that the T2 TIC from one compartment plays a dominant role in OA progression, we isolated the medial and lateral sub-populations from the dominant compartment and compared the compartment to the mechanical axis from standing full-length limb radiographs. Individuals with a dominant compartment on the lateral condyle were highly correlated with valgus alignment, and individuals associated with a dominant compartment on the medial condyle were associated with a varus alignment. A comparison of these two populations demonstrated the differences were statistically significant as measured by the Student *t*-test; however the notched box plot shows the 95% confidence interval of the means' overlap, suggesting that, at a minimum, the dominant compartment's location is highly correlated with mechanical axis [Fig. 4(C)].

Discussion

Measurement of cartilage T2 heterogeneity using the T2 TIC represents a possible biomarker for symptomatic OA. Interestingly, in a majority of individuals, the main contribution of the T2 TIC originated from one dominant compartment which is highly correlated with the mechanical alignment suggesting the measured differences in cartilage T2 between these groups are associated with alteration in joint loading. This suggests that for most subjects, the T2 TIC in a single knee compartment is predicting the onset of OA symptoms,

and this compartment is correlated with areas of increased joint loading based on the mechanical alignment of the knee.

The T2 TIC is a composite measure of the inherent signal variation of the T2 map. It represents a series of texture metrics that capture the loss of the normal signal pattern. For example, GLRL and GLCM texture features measure this change in signal homogeneity by measuring the repetition of voxel signal intensity across an entire voxel neighborhood. Normal knee articular cartilage T2 values have a spatial signal variation in T2 values that reflect regional variation in the collagen fiber anisotropy and water content⁴. In early OA, there is a loss of cartilage anisotropy and increase in water content that produce greater variability in T2 values between neighboring voxels¹⁰. The T2 TIC measures the loss of the articular pattern of signal variation between the articular surface and bone observed in normal cartilage. A low TIC represents a homogenous signal and a high TIC represents a heterogeneous signal. From a structural perspective, it is likely that the T2 TIC is driven by multiple mechanisms of cartilage degradation that occur concurrently with the onset of OA. These include processes known to lead to elevation in cartilage T2 such as loss of collagen anisotropy and increased cartilage water content^{4,7,29}, as well as other factors that can lead to greater heterogeneity in water bindings sites on macromolecules contributing to T2 relaxation. For example, the absence of aggrecan in cartilage has been shown to produce high variability in the T2 distribution of cartilage³⁰.

The importance of texture in T2 relaxation time mapping has been recently demonstrated by multiple groups. The value of the

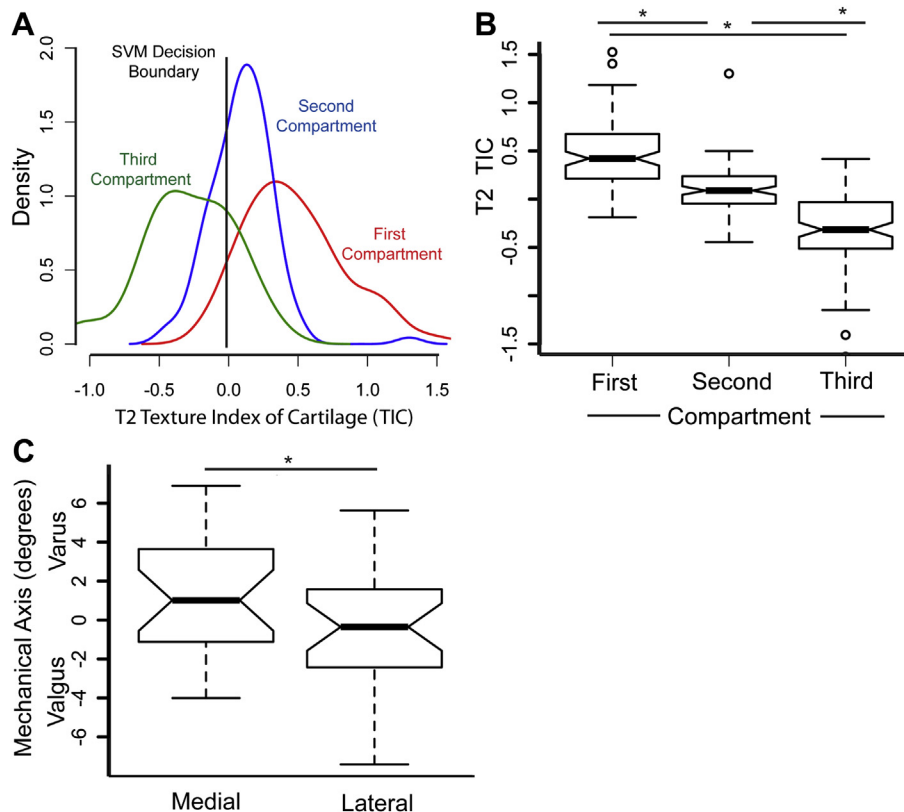


Fig. 4. A T2 TIC that indicates OA is associated primarily with one knee compartment. (A) The features that dominate OA decisions, for most subjects, come from primarily one knee compartment (medial, lateral, or patella). To demonstrate this, the aggregate TIC “partial scores” were calculated for each knee compartment in each subject. A histogram is shown for the compartment in rank order with the largest partial score (first), for the second largest partial score (second), and for the minimum partial score (third), across subjects. The SVM decision score is shown as the vertical black line. (B) The average TIC for each of the three compartments (First, Second, and Third) are well-separated. Notched box plots show the average TIC of each of the compartments. (C) The dominant compartment that predicted OA progression is strongly correlated with mechanical alignment. Notched box plot of individuals with a dominant medial or lateral compartment contribution to the TIC compared as a function of the mechanical axis. Individuals with a varus alignment were associated with an increased TIC in the medial compartment and vice versa for valgus alignment. Negative mechanical axis values indicate valgus alignment. **P* < 0.05.

GLCM texture metric changes on a longitudinal basis in a small sub-population of individuals²³. When comparing unexposed control and populations at increased risk of developing OA, the mean T2 signal, GLCM contrast, and variance are elevated in the at risk group³¹. In populations that have already developed clinically significant OA, differences in signal texture have been measured as compared to unexposed control populations²⁴. Our results support these findings and extend this work: First, changes in texture are not only present prior to the development of clinical or radiographic knee OA but are also prognostic for symptomatic OA progression. Second, these signal changes are localized to a single compartment, which is correlated with the mechanical axis. Together, these results support further development of the T2 TIC as a possible prognostic imaging biomarker to identify individuals at risk for symptomatic OA progression and to explore the association with knee mechanics.

The non-progression and symptomatic OA progression populations were similar at baseline from a clinical perspective. Asymptomatic OA was defined as a WOMAC score less than 10 and symptomatic OA progression was defined as a change of at least 10. The minimum change in the WOMAC subscales has been measured between 9 and 20^{32–34}. In patients with low baseline WOMAC scores, the minimally perceptible change was approximately 10³². Here, the differences in the baseline WOMAC scores between the two groups are similar from a clinical perspective.

The features used to build the T2 TIC are not a unique solution. MFE was used to reduce the number of features from 725 to 22 by eliminating redundant features or ones that did not add value to the model. Different features eliminated from the model can be substituted for features included in the model without a loss of accuracy. We reported the maximum accuracy we could achieve from our data set. The description of one feature set reveals that a number of features across a range of sections are important.

To simplify the model, the tibia was not segmented based on the assumption that T2 signal changes on the femur would reflect corresponding changes on the tibia to at least a minimal degree. Wear patterns of the tibiofemoral articulation are known to correspond between surfaces³⁵. Other groups have demonstrated corresponding contact pressure and contact areas in the tibiofemoral articular surface in both compartments despite alterations in varus or valgus loading³⁶. Addition of the tibia articular surface into the model would likely improve the accuracy of the prognostic biomarker.

To increase the cohort size and accuracy of our model, patients with a low level of radiographic OA (KL grade 2) were included in the cohort. Although statistically significant, the difference in the mean KL grades between the two populations were less than one and significantly below the threshold defined by the grading scale for definite OA of KL ≥ 2 . If the cohorts excluded subjects with an initial KL grade of 2 from the analysis, there was only a minimal loss of accuracy of the model, suggesting the T2 TIC was primarily prognostic for changes in WOMAC score.

The OAI has demonstrated consistent and reproducible T2 map values across multiple sites and time periods^{37,38}. Variation in T2 values has been shown to be less than approximately 5% in this group³⁷. Further, the signal on each T2-weighted voxel decays exponentially. This exponential decay can be used to transform a T2-weighted image to a T2 map that possesses the information of T2-weighted image in a format that is independent of user and machine characteristics increasing the reproducibility of T2 mapped values⁴. In this work, we demonstrated our model's ability to handle T2 maps from multiple centers by using T2 maps collected from multiple sites at different time points.

An advantage of using texture analysis rather than absolute T2 values is less sensitivity to systematic bias between scanners in the T2 measurement. A recent study has shown that T2 measurements

between different vendors can vary by as much as 5 ms–10 ms making it difficult to compare T2 values across different scanners³⁹. Since texture analysis is comparing differences in T2 values between neighboring voxels it is insensitive to factors that produce a uniform systematic bias in the measurement. Factors such as the magic angle and partial volume effect produce regional variation in T2 values across cartilage that makes it difficult to compare T2 values obtained from different regions in the joint. Since these artifacts occur in a very predictable pattern within the joint, analysis of T2 texture is able to differentiate this pattern from the more random spatial distribution in T2 seen with early cartilage injury.

In conclusion, our results indicate that measurement of cartilage T2 heterogeneity as defined by a composite marker termed the T2 TIC is able to differentiate subjects with preclinical OA that are at risk of developing OA symptoms. The ability to differentiate patients at risk for symptomatic OA progression prior to symptomatic presentation based on MRI signal changes would be valuable in clinical and epidemiological studies for disease-modifying OA drugs (DMOADs), joint preservation surgical interventions, and the development of post-traumatic arthritis in anterior cruciate ligament and meniscal injuries. Further studies are needed using the T2 TIC in populations outside of the OAI to validate accuracy.

Author contributions

All of the authors were involved in experimental or software design, data analysis, and manuscript preparation.

Conflict of interest

The authors have no conflicts of interest to declare.

Acknowledgments

Funding support was provided in part by the Woodward Family Endowment in Biomedical Engineering to the Penn State University. The OAI is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, and conducted by the OAI Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. Private sector funding for the OAI is managed by the Foundation for the National Institutes of Health. This manuscript was prepared using an OAI public use data set and does not necessarily reflect the opinions or views of the OAI investigators, the NIH, or the private funding partners.

Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.joca.2013.06.007>.

References

1. Brandt KD, Fife RS, Braunstein EM, Katz B. Radiographic grading of the severity of knee osteoarthritis: relation of the Kellgren and Lawrence grade to a grade based on joint space narrowing, and correlation with arthroscopic evidence of articular cartilage degeneration. *Arthritis Rheum* 1991;34:1381–6.
2. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskelet Disord* 2008;9:116.
3. David-Vaudey E, Ghosh S, Ries M, Majumdar S. T2 relaxation time measurements in osteoarthritis. *Magn Reson Imaging* 2004;22:673–82.

4. Mosher TJ, Dardzinski BJ. Cartilage MRI T2 relaxation time mapping: overview and applications. *Semin Musculoskelet Radiol* 2004;8:355–68.
5. Mosher TJ, Smith H, Dardzinski BJ, Schmithorst VJ, Smith MB. MR imaging and T2 mapping of femoral cartilage: in vivo determination of the magic angle effect. *AJR Am J Roentgenol* 2001;177:665–9.
6. Wheaton AJ, Casey FL, Gougoutas AJ, Dodge GR, Borthakur A, Lonner JH, et al. Correlation of T1rho with fixed charge density in cartilage. *J Magn Reson Imaging* 2004;20:519–25.
7. Burstein D, Gray ML. Is MRI fulfilling its promise for molecular imaging of cartilage in arthritis? *Osteoarthritis Cartilage* 2006;14:1087–90.
8. Xia Y, Moody JB, Alhadlaq H. Orientational dependence of T2 relaxation in articular cartilage: a microscopic MRI (microMRI) study. *Magn Reson Med* 2002;48:460–9.
9. Harrison MM, Cooke TD, Fisher SB, Griffin MP. Patterns of knee arthrosis and patellar subluxation. *Clin Orthop Relat Res* 1994; 56–63.
10. Yulish BS, Lieberman JM, Newman AJ, Bryan PJ, Mulopulos GP, Modic MT. Juvenile rheumatoid arthritis: assessment with MR imaging. *Radiology* 1987;165:149–52.
11. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833–40.
12. Nevitt MC, Felson DT, Lester G. The Osteoarthritis Initiative: Protocol for the Cohort Study. UC San Francisco; Boston University; National Institute of Arthritis, Musculoskeletal Skin Diseases, <http://oai.epi-ucsf.org/datarelease/docs/StudyDesignProtocol.pdf>.
13. 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;16:1433–41.
14. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. *Ann Rheum Dis* 1957;16:494–502.
15. McGrory JE, Trousdale RT, Pagnano MW, Nigbur M. Preoperative hip to ankle radiographs in total knee arthroplasty. *Clin Orthop Relat Res* 2002;196–202.
16. Mattes D, Haynor DR, Vesselle H, Lewellen TK, Eubank W. PET-CT image registration in the chest using free-form deformations. *IEEE Trans Med Imaging* 2003;22:120–8.
17. Urish KL, Durkin JR, Williams A, Chu CR. Registration of magnetic resonance image series for knee articular cartilage analysis: data from the osteoarthritis initiative. *Cartilage* 2013;4:20–7.
18. Yoo TS, Ackerman MJ, Lorensen WE, Schroeder W, Chalana V, Aylward S, et al. Engineering and Algorithm Design for an Image Processing API: A Technical Report on ITK – the Insight Toolkit. Amsterdam: IOS Press; 2002.
19. Durkin JR, Miller DJ, Urish KL. Creating a 2D Active Shape Model Using itk::ImagePCAShapeModelEstimator. *IJ*; 2011. <http://hdl.handle.net/10380/3282>.
20. 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;14:50–5.
21. Qazi AA, Dam EB, Nielsen M, Karsdal MA, Pettersen PC, Christiansen C. Osteoarthritic cartilage is more homogeneous than healthy cartilage: identification of a superior region of interest colocalized with a major risk factor for osteoarthritis. *Acad Radiol* 2007;14:1209–20.
22. Qazi AA, Folkesson J, Pettersen PC, Karsdal MA, Christiansen C, Dam EB. Separation of healthy and early osteoarthritis by automatic quantification of cartilage homogeneity. *Osteoarthritis Cartilage* 2007;15:1199–206.
23. Carballido-Gamio J, Link TM, Majumdar S. New techniques for cartilage magnetic resonance imaging relaxation time analysis: texture analysis of flattened cartilage and localized intra- and inter-subject comparisons. *Magn Reson Med* 2008;59: 1472–7.
24. Blumenkrantz G, Stahl R, Carballido-Gamio J, Zhao S, Lu Y, Munoz T, et al. The feasibility of characterizing the spatial distribution of cartilage T(2) using texture analysis. *Osteoarthritis Cartilage* 2008;16:584–90.
25. Boniatis I, Costaridou L, Cavouras D, Kalatzis I, Panagiotopoulos E, Panayiotakis G. Assessing hip osteoarthritis severity utilizing a probabilistic neural network based classification scheme. *Med Eng Phys* 2007;29:227–37.
26. Guyon I, Weston J, Barnhill S. Gene selection for cancer classification using support vector machines. *Machine Learn* 2002;46:389–422.
27. Aksu Y, Miller DJ, Kesidis G, Yang Q. Margin-maximizing feature elimination methods for linear and nonlinear kernel support vector machines. *IEEE Trans Neural Networks* 2010;21:701–17.
28. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCr: visualizing classifier performance in R. *Bioinformatics* 2005;21:3940–1.
29. Hoch DH, Grodzinsky AJ, Koob TJ, Albert ML, Eyre DR. Early changes in material properties of rabbit articular cartilage after meniscectomy. *J Orthop Res* 1983;1:4–12.
30. Mosher TJ, Chen Q, Smith MB. 1H magnetic resonance spectroscopy of nanomelic chicken cartilage: effect of aggrecan depletion on cartilage T2. *Osteoarthritis Cartilage* 2003;11: 709–15.
31. Joseph GB, Baum T, Carballido-Gamio J, Nardo L, Virayavanich W, Alizai H, et al. Texture analysis of cartilage T2 maps: individuals with risk factors for OA have higher and more heterogeneous knee cartilage MRT2 compared to normal controls - data from the osteoarthritis initiative. *Arthritis Res Ther*; 13: R153.
32. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. *Ann Rheum Dis* 2005;64: 29–33.
33. Ehrlich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. *J Rheumatol* 2000;27: 2635–41.
34. Bellamy N, Carette S, Ford PM, Kean WF, le Riche NG, Lussier A, et al. Osteoarthritis antirheumatic drug trials. III. Setting the delta for clinical trials—results of a consensus development (Delphi) exercise. *J Rheumatol* 1992;19:451–7.
35. Raju PK, Kini SG, Verma A. Wear patterns of tibiofemoral articulation in osteoarthritic knees: analysis and review of literature. *Arch Orthop Trauma Surg* 2012;132:1267–71.
36. Riegger-Krugh C, Gerhart TN, Powers WR, Hayes WC. Tibiofemoral contact pressures in degenerative joint disease. *Clin Orthop Relat Res* 1998;233–45.
37. Schneider E, Nessesaver M. The Osteoarthritis Initiative (OAI) magnetic resonance imaging quality assurance update. *Osteoarthritis Cartilage* 2013;21:110–6.
38. Schneider E, Nessesaver M, White D, Purdy D, Martin L, Fanella L, et al. The osteoarthritis initiative (OAI) magnetic resonance imaging quality assurance methods and results. *Osteoarthritis Cartilage* 2008;16:994–1004.

39. Balamoody S, Williams TG, Wolstenholme C, Waterton JC, Bowes M, Hodgson R, et al. Magnetic resonance transverse relaxation time T2 of knee cartilage in osteoarthritis at 3-T: a cross-sectional multicentre, multi-vendor reproducibility study. *Skeletal Radiol* 2013;42: 511–20.
40. Keffalas MG. Feature selection and classification to automatically detect knee osteoarthritis using MRI. In: Department of Electrical Engineering. Univeristy Park, PA: The Pennsylvania State University; 2010:45.
41. Chang C-C, Lin C- J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2011;2:1–27.
42. Guyon I, André Elisseeff. An introduction to variable and feature selection. *J Mach Learn Res* 2003;3:1157–82.
43. Aksu Y, Miller DJ, Kesidis G, Yang QX. Margin-maximizing feature elimination methods for linear and nonlinear Kernel-based discriminant functions. *Neural Networks. IEEE Trans* 2010;21:701–17.
44. Kohavi R, John GH. Wrappers for feature subset selection. *Artif Intell* 1997;97:273–324.