Osteoarthritis and Cartilage 25 (2017) 1107–1113

Osteoarthritis and Cartilage

Quantitative measurement of medial femoral knee cartilage volume – analysis of the OA Biomarkers Consortium FNIH Study cohort

L.F. Schaefer † *, M. Sury †, M. Yin †, S. Jamieson †, I. Donnell †, S.E. Smith †, J.A. Lynch ‡, M.C. Nevitt ‡, J. Duryea †

† *Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA*
‡ *UCSF, San Francisco, CA, USA*

ARTICLE INFO

Article history:
Received 13 June 2016
Accepted 22 January 2017

Keywords:
Osteoarthritis
Cartilage
Knee
Magnetic resonance imaging
Segmentation software

SUMMARY

Objective: Large studies of knee osteoarthritis (KOA) require well-characterized efficient methods to assess progression. We previously developed the local-area cartilage segmentation (LACS) software method, to measure cartilage volume on magnetic resonance imaging (MRI) scans. The present study further validates this method in a larger patient cohort and assesses predictive validity in a case–control study.

Method: The OA Biomarkers Consortium FNIH Project, a case–control study of KOA progression nested within the Osteoarthritis Initiative (OAI), includes 600 subjects in four subgroups based on radiographic and pain progression. Our software tool measured change in medial femoral cartilage volume in a central weight-bearing region. Different sized regions of cartilage were assessed to explore their sensitivity to change. The readings were performed on MRI scans at the baseline and 24-month visits. We used standardized response means (SRMs) for responsiveness and logistic regression for predictive validity.

Results: Cartilage volume change was associated strongly with radiographic progression (odds ratios (OR) = 4.66; 95% confidence intervals (CI) = 2.85–7.62). OR were significant but of lesser magnitude for the combined radiographic and pain progression outcome (OR = 1.70; 95% CI = 1.40–2.07). For the full 600 subjects, theSRM was −0.51 for the largest segmented area. Smaller areas of cartilage segmentation were also able to predict the case–control status. The average reader time for the largest area was less than 20 min per scan. Smaller areas could be assessed with less reader time.

Conclusion: We demonstrated that the LACS method is fast, responsive, and associated with radiographic and pain progression, and is appropriate for existing and future large studies of KOA.

Introduction

Osteoarthritis (OA) is a highly prevalent, disabling disease, with tremendous individual and socioeconomic burden that mostly affects elderly people[1]. To date there are no disease modifying treatment options available for OA.

Imaging is important for the diagnosis and assessment of OA both in the clinical setting as well as the research environment. For knee OA, radiography continues to be a primary imaging modality, due to convenience and cost. However, soft tissue such as cartilage cannot be visualized directly in knee radiographs; joint space width provides only an indirect measurement of cartilage. Although more costly and time-consuming than radiography, magnetic resonance imaging (MRI), with its good soft tissue contrast, is the superior option for evaluating some OA-related structures such as cartilage, bone marrow lesions, and the menisci[2].

Objective, reliable, and fast methods to determine knee cartilage volume are needed for large OA trials and observational research. Existing studies of knee OA such as the Osteoarthritis Initiative (OAI)[3,4] and the Multicenter Osteoarthritis Study (MOST)[5] each have tens of thousands of individual knee MRI scans. The vast

* Address correspondence and reprint requests to: Lena F. Schaefer, Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA.
E-mail address: lenafranziskaschaefer@yahoo.com (L.F. Schaefer).

http://dx.doi.org/10.1016/j.joca.2017.01.010
1063-4584/

majority of the MRI data from the OAI has not been assessed for quantitative cartilage measures since resources were not provided for this purpose. Additionally, large numbers of MRI data sets will likely be necessary to appropriately power future longitudinal clinical trials of knee OA. Clinical trials for OA therapies also rely on highly responsive measures of structural change so that statistically significant differences in disease progression between the treatment and placebo arms can be found. Quantitative as well as semi-quantitative methods to assess cartilage status in knee OA exist[2]. Semi-quantitative methods[6,7] are based on a qualitative assessment and provide ordinal rather than continuous measurements, and can be cost intensive due to high reading time and the requirement for an experienced reader with specialized training and expertise in musculoskeletal radiology. Quantitative methods were reported to be superior to semi-quantitative methods in assessing cartilage change for knee OA[8,9]. Efforts to decrease MRI reader time, while maintaining performance, directly address the high cost of radiological imaging for current and future studies of knee OA.

We previously developed and validated an efficient, reproducible, and responsive quantitative software tool to measure cartilage volume in focal locations on the medial femur on MRI scans[10,11]. The local-area cartilage segmentation (LACS) method uses anatomical landmarks and a mathematically robust coordinate system to identify consistent regions of cartilage for fast segmentation. The goal of our current study is to further explore the responsiveness and examine clinical validity in a substantially larger cohort by applying LACS to a case–control sample of knee OA progression, and investigate refinements to improve efficiency. We expect to demonstrate that this method is ideal for existing and future large studies of knee OA that use MRI.

Methods

Study design and cohort

For this study, we analyzed subjects that make up the OA Biomarkers Consortium FNIH Study (https://oai.epi-ucsf.org/datarelease/FNIH.asp), a nested case–control study within the OAI. The OAI is a longitudinal cohort study of 4796 men and women ages 45–79 with, or at risk for, knee OA at the beginning of the study. Knee radiography and MRI as well as a clinical assessment were performed annually. In addition, biochemical specimens were collected from all participants. A primary objective of the OAI is to create a public resource for identifying, characterizing, and validating a broad range of imaging biomarkers for OA of the knee that could be used to investigate basic research hypotheses and to serve as outcomes in clinical trials of new therapies[4].

The goal of the OA Biomarkers Consortium FNIH Study is to find structural and biochemical biomarkers for radiographic and pain progression in knees with mild to moderate OA. Details of the design of this study have been published elsewhere[12–14]. Briefly, it includes 600 subjects in four subgroups based on radiographic and pain progression. Radiographic progression was defined by medial tibiofemoral joint space loss $\geq$0.7 mm from baseline to 24, 36, or 48 months measured on conventional radiographs acquired with a fixed-flexion protocol[15]. Pain progression was defined as a persistent increase on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain scale between 24 and 60 months. Based on this, the cohort was subdivided into four groups:

Group 1: radiographic and pain progressors ($n = 194$).
Group 2: radiographic-only progressors ($n = 103$).
Group 3: pain-only progressors ($n = 103$).
Group 4: no radiographic or pain progressors ($n = 200$).

For the main analysis, we adhere to the case/control definition established by the FNIH researchers[16] having both radiographic and pain progression (Group 1) as cases and combining the other three groups (2, 3, and 4) to be controls. For secondary analyses, and to make a direct comparison with another study[17], we investigated comparisons apart from the main analysis by individually comparing Groups 1, 2, and 3 with Group 4, and all subjects with radiographic progression (Group 1 and 2) to subjects without radiographic progression (Groups 3 and 4). Similarly we combined all patients with pain progression (Groups 1 and 3) and compared them to all patients without pain progression (Groups 2 and 4).

Image analysis and reader procedure

The readings were performed on sagittal double echo steady state (DESS) 3D MRI scans (sagittal, 0.365 mm 0.365 mm, 0.7-mm slice thickness, repetition time 16.5 ms, echo time 4.7 ms) at the baseline and 24-month visits with the reader (LS) blinded to time point and case–control status (see Fig. 1). All knees were evaluated at a fixed measurement location in the central weight-bearing portion of the medial femur as described in a previous publication[10]. Briefly, the measurement region was based on two axes of a cylindrical coordinate system, z and θ. z roughly corresponded to the medial-lateral direction (greater z was more medial) and θ to the anterior-posterior part of the articular surface of the femur (greater θ was more posterior). Since the coordinate system is linked to anatomical landmarks, the effective size (in mm) of the z variable changes with knee size.

Using custom software, the reader used the LACS method to segment cartilage in the regions specified by the coordinate system. The software informed the reader of the slices required to evaluate the cartilage in the required z-range, and the limits on each slice to ensure coverage in θ. Automated image analysis tools were also provided to increase speed and objectivity, including edge detection algorithms that the reader could initiate in areas adjacent to the cartilage margins and a method for the reader to indicate areas of denuded cartilage. The automated steps increased objectivity and minimized the need for manual segmentation by providing tools to allow the reader to guide the automated software when corrections were required.

Reproducibility

The LACS method was previously validated for intra-reader and inter-reader reliability and showed good reproducibility with intraclass correlation coefficients (ICCs) > 0.9)[10]. In a recent published study the repositioning error of the LACS method for $\Delta z = 0.1$ was measured on 10 healthy volunteers with a modified DESS-sequence twice on the same day where the subjects were

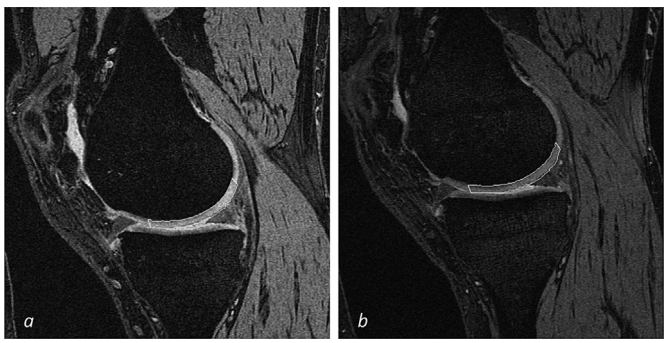

Fig. 1. Corresponding segmented cartilage areas at baseline (a) and follow-up (b).

removed and placed back into the MRI scanner in-between scans. The coefficient of variation (CoV) was 3.2% for repositioning precision and the ICC was >0.9[11].

In addition, a random subset of 20 FNIH subjects was chosen to evaluate intra and inter-reader reproducibility for our current study. The baseline and 24-month scans were read a second time by the primary reader (LS) and a second reader (ID) with the procedure described above. Since the previously published data[11] were from healthy volunteers and used a different DESS sequence than the one acquired in the OAI, we conducted a new study to investigate repositioning reproducibility. An independent set of DESS scans from 15 participants was used to assess repositioning reproducibility and were assessed by the reader (LS) in a fully blinded manner. These participants were from an OAI pilot study of the reproducibility of the OAI MRI sequences whose OA and other characteristics were similar to those of OAI participants[18]. The subjects were scanned, removed from the magnet, walked for 10 min, and scanned a second time. These data were also used in a separate publication to assess repositioning reproducibility for a full-plate cartilage volume method[17].

We chose to evaluate the method at a location approximately in the central weight bearing region of the medial femur centered at $z = 0.8$ and $\theta = 210°$, which demonstrated substantial cartilage change in two previous studies[10,19]. The size of the region in θ, ($\Delta\theta$) was 100° and 0.09 in z (Δz). For a secondary analysis and to assess potential reduction to reader speed, we varied the range in z (Δz) from 0.09 to 0.01 representing increasingly smaller areas of cartilage, each centered on the same point. Varying Δz alone (not $\Delta\theta$) corresponded directly to a different numbers of slices requiring reader attention, and therefore the assessment was potentially more efficient than with varying $\Delta\theta$. Fig. 2 shows 3D renderings of the segmented femur. Fig. 2(a) represents the largest segmentation area ($\Delta z = 0.09$). Fig. 2(b)–(d) illustrates smaller segmentation areas corresponding to reduced values of Δz, which could be assessed using existing segmentations. Once the optimal region is determined, future studies could employ smaller regions further reducing the reader time since fewer slices would require attention.

Statistical analysis

The primary goal of the study was to differentiate between the case–control status of the study cohort and assess responsiveness. We used logistic regression with odds ratios (OR) and 95% confidence intervals (CI) as metrics to assess the association of change in cartilage volume with progression status. Each model was adjusted for age, sex, body mass index (BMI), and race. ORs were calculated per 1 standard deviation (SD) loss of cartilage from baseline to 24 months.

Based on the assumption that knee OA worsens structurally over time, the measurement of change in cartilage volume reflects the performance of imaging-based outcome measures. Responsiveness was measured by calculating the average and SD of the volume change and the standardized response mean (SRM) (average ΔV/SD of ΔV). Precision was assessed using ICCs and the CoV.

Results

The demographic characteristics for the 600 subjects are given in Table I.

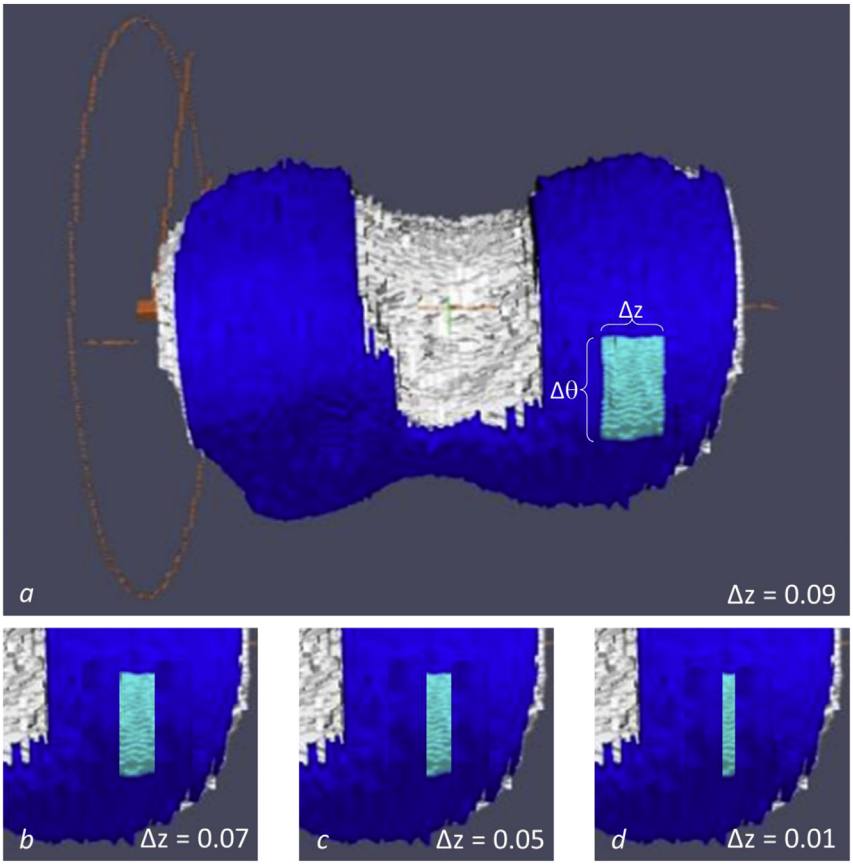

Fig. 2. 3D rendering of the segmented femur showing examples of different sized regions of cartilage segmentation in light blue.

Table I
Demographics of the study population

	Cases Group 1 $n = 194$	Controls Group 2, 3, 4 $n = 406$	Group 2 $n = 103$	Group 3 $n = 103$	Group 4 $n = 200$
Age (years)	62.0 ± 8.8	61.3 ± 8.9	63.1 ± 8.3	59.2 ± 8.7	61.5 ± 9.1
Female sex (%)	56.7	59.9	44.7	65.0	65.0
BMI (kg/m^2)	30.7 ± 4.8	30.7 ± 4.8	30.7 ± 4.7	31.1 ± 5.0	30.5 ± 4.8
White (%)	80.0	78.8	88.3	71.8	77.5
History of injury to index knee (%)	35.2	35.8	39.8	37.0	33.2
KLG 1/2/3 at baseline, (n, %)	24 (12.4)/84 (43.3)/86 (44.3)	51 (12.6)/222 (54.7)/133 (32.8)	14 (13.6)/47 (45.6)/42 (40.8)	13 (12.6)/61 (59.2)/29 (28.1)	24 (12.0)/114 (57.0)/62 (31.0)
Baseline WOMAC pain score	10.2 ± 13.0	13.0 ± 16.7	16.5 ± 19.9	9.6 ± 13.3	13.0 ± 16.2
Baseline minimum medial radiographic JSW (mm)	3.8 ± 1.4	3.9 ± 1.1	3.8 ± 1.2	3.9 ± 1.0	3.9 ± 1.0

Except where indicated otherwise, values are the mean ± SD.
KLG = Kellgren–Lawrence grade.
JSW = joint space width.

The OR for the case/control analysis are given in Table II. For the primary case/control analysis (Group 1 vs Groups 2, 3, and 4), the ORs were significant. The largest area with a Δz of 0.09 provided an OR of 1.70 with 95% CI of 1.40–2.07. The highest OR in this group was seen for $\Delta z = 0.04$ (OR = 1.84; 95% CI = 1.50–2.25). The highest ORs were seen for the comparison between knees with radiographic progression vs knees with neither radiographic nor pain progression (Group 2 vs Group 4). In this group, the highest OR was seen for a $\Delta z = 0.07$ (OR = 4.69; 95% CI = 2.85–7.72). The largest area with a Δz of 0.09 provided an OR of 4.66 with 95% CI of 2.85–7.62. Similar results were seen in the analysis of all subjects with radiographic progression combined vs all subjects without radiographic progression (Group 1 + 2 vs Group 3 + 4) ($\Delta z = 0.09$; OR = 3.13; 95% CI = 2.37–4.12). Adjusting for semi-quantitative meniscus and bone marrow lesion scores had no substantial impact on the OR. For the analysis of knees with pain vs knees without pain or radiographic progression (Group 3 vs Group 4) only one borderline significant result was observed with the smallest Δz (OR = 1.41; 95% CI = 1.00–1.99). The largest area with a Δz of 0.09 provided an OR of 1.08 with 95% CI of 0.69–1.71. Similarly in the analysis of all subjects with pain progression combined vs all subjects without pain progression (Group 1 + 3 vs Group 2 + 4), only borderline significant results were observed ($\Delta z = 0.09$; OR = 1.26; 95% CI = 1.06–1.51).

Responsiveness is given in Table III for each group and Δz value. For the full 600 subjects, the highest SRM in magnitude was at $\Delta z = 0.07$ and 0.08 (SRM of −0.52) and further reduction of the segmented area reduced the SRM only slightly until $\Delta z = 0.02$ (SRM −0.46). The largest segmented area with a Δz of 0.09 provided a similar SRM of −0.51. For clinical trials, the power of a study is directly proportional to the SRM squared. This implies, for example, that using $\Delta z = 0.03$ (SRM = 0.48) vs $\Delta z = 0.09$ (SRM = 0.51) would require only 12% more subjects to achieve the same power, but would reduce the reader time by 67%.

The precision results are given in Table IV. Intra-reader was similar to the repositioning reproducibility and generally better than inter-reader reproducibility. As with the ORs and responsiveness, the precision was substantially worse for very low values of Δz.

The total average reader time for the largest area of cartilage segmentation was less than 20 min per scan, split approximately evenly between a research assistant and a more skilled reader. There was, in general, little difference in the ORs (Table II) or SRM values (Table III) for Δz above 0.02. Together, these data suggest that using substantially smaller sub regions has minimal impact and that the readings could be done even more efficiently without substantial loss of performance.

Discussion

We have established clinical validity of the LACS method in a case–control setting. These results indicate that the method is

Table II
Baseline to 24 month change in cartilage volume*

Δz	Primary analysis: cases vs controls†	Group 1 vs 4†	Group 2 vs 4†	Group 3 vs 4†		Group 1, 2, 3 vs 4†	Groups 1 + 2 vs 3 + 4†	Groups 1 + 3 vs 2 + 4†	
	Adjusted OR (95% CI)	Adjusted OR (95% CI)	Adjusted OR (95% CI)	Adjusted OR (95% CI)	P-value	Adjusted OR (95% CI)	Adjusted OR (95% CI)	Adjusted OR (95% CI)	P-value
0.01	1.58 (1.31–1.90)‡	2.30 (1.73–3.06)‡	2.61 (1.84–3.69)*	1.41 (1.00–1.99)	0.0489	2.02 (1.60–2.56)‡	2.12 (1.70–2.65)‡	1.34 (1.12–1.60)	0.0012
0.02	1.70 (1.40–2.06)‡	3.00 (2.14–4.20)‡	3.66 (2.35–5.72)‡	1.10 (0.72–1.66)	0.6611	2.21 (1.71–2.86)‡	2.90 (2.22–3.78)‡	1.28 (1.07–1.53)	0.0063
0.03	1.82 (1.49–2.22)‡	3.48 (2.43–5.01)‡	3.72 (2.38–5.80)‡	1.02 (0.66–1.58)	0.9359	2.34 (1.79–3.06)‡	3.19 (2.41–4.22)‡	1.32 (1.11–1.58)	0.0022
0.04	1.84 (1.50–2.25)‡	3.75 (2.57–5.45)‡	4.47 (2.75–7.25)‡	1.13 (0.71–1.77)	0.6109	2.56 (1.93–3.39)‡	3.44 (2.57–4.60)‡	1.35 (1.13–1.62)	0.0012
0.05	1.81 (1.48–2.21)‡	3.56 (2.47–5.12)‡	4.19 (2.63–6.68)‡	1.11 (0.71–1.75)	0.6459	2.49 (1.89–3.29)‡	3.35 (2.52–4.45)‡	1.33 (1.11–1.59)	0.0018
0.06	1.82 (1.49–2.23)‡	3.58 (2.49–5.14)‡	4.50 (2.78–7.29)‡	1.15 (0.74–1.80)	0.5371	2.53 (1.92–3.34)‡	3.38 (2.54–4.49)‡	1.34 (1.12–1.60)	0.0015
0.07	1.79 (1.47–2.18)‡	3.32 (2.33–4.73)‡	4.69 (2.85–7.72)‡	1.08 (0.68–1.71)	0.7374	2.45 (1.86–3.23)‡	3.28 (2.47–4.36)‡	1.31 (1.10–1.57)	0.0029
0.08	1.73 (1.42–2.11)‡	3.08 (2.19–4.33)‡	4.34 (2.69–7.00)‡	1.09 (0.69–1.71)	0.7202	2.33 (1.78–3.05)‡	3.09 (2.35–4.07)‡	1.29 (1.08–1.54)	0.0058
0.09	1.70 (1.40–2.07)‡	3.04 (2.17–4.28)‡	4.66 (2.85–7.62)‡	1.08 (0.69–1.71)	0.7372	2.32 (1.78–3.03)‡	3.13 (2.37–4.12)‡	1.26 (1.06–1.51)	0.0107

The highest OR for each group is highlighted in red.
Group 1 was defined as cases in the primary analysis. Group 2, 3 and 4 were combined as controls in the primary analysis.
* ORs with 95 % CIs and P-values are provided as a function of Δz and $\Delta\theta$. The center point is located at $z_0 = 0.8$, $\theta_0 = 210°$.
† Group 1, consisted of knees with both radiographic and pain progression (primary cases, $n = 194$); Group 2, consisted of knees with radiographic progression only ($n = 103$); Group 3, consisted of knees with pain progression only ($n = 103$); Group 4, consisted of knees with neither radiographic nor pain progression ($n = 200$).
‡ P-value < 0.0001.

Table III
Responsiveness to change over 24 months*

Δz	All $n = 600$	Radiographic and pain progression (Group 1) $n = 194$	Radiographic progression only (Group 2) $n = 103$	Pain progression only (Group 3) $n = 103$	Neither radiographic nor pain progression (Group 4) $n = 200$	Controls (Group 2, 3, and 4) $n = 406$
	% Change (SRM)	% Change (SRM)	% Change (SRM)	% Change (SRM)	% Change (SRM)	% Change (SRM)
0.01	−7.0 (−0.42)	−12.9 (−0.61)	−12.5 (−0.71)	−3.9 (−0.32)	−0.9 (−0.08)	−4.5 (−0.32)
0.02	−6.5 (−0.46)	−12.2 (−0.68)	−13.0 (−0.77)	−1.6 (−0.17)	−1.1 (−0.14)	−4.1 (−0.34)
0.03	−6.6 (−0.48)	−12.7 (−0.73)	−12.4 (−0.76)	−1.5 (−0.19)	−1.2 (−0.18)	−4.0 (−0.36)
0.04	−6.5 (−0.49)	−12.6 (−0.73)	−12.2 (−0.79)	−1.7 (−0.22)	−1.0 (−0.16)	−3.9 (−0.37)
0.05	−6.5 (−0.51)	−12.2 (−0.74)	−12.1 (−0.81)	−1.7 (−0.24)	−1.2 (−0.19)	−4.0 (−0.39)
0.06	−6.2 (−0.51)	−11.8 (−0.74)	−11.7 (−0.83)	−1.7 (−0.23)	−1.1 (−0.18)	−3.9 (−0.39)
0.07	−6.2 (−0.52)	−11.5 (−0.73)	−11.7 (−0.84)	−1.7 (−0.25)	−1.3 (−0.22)	−4.0 (−0.41)
0.08	−6.1 (−0.52)	−11.0 (−0.72)	−11.5 (−0.84)	−1.8 (−0.27)	−1.4 (−0.24)	−4.0 (−0.42)
0.09	−5.9 (−0.51)	−10.5 (−0.71)	−11.5 (−0.86)	−1.6 (−0.25)	−1.4 (−0.23)	−3.9 (−0.42)

* The % of cartilage change as well as SRMs values are provided as a function of Δz. The center point is located at $z_0 = 0.8$, $\theta_0 = 210^\circ$ and $\Delta\theta = 100^\circ$.

Table IV
Measurements of precision*

Δz	Intra-reader reproducibility $n = 40$		Inter-reader reproducibility $n = 40$		Repositioning reproducibility $n = 15$	
	ICC	CoV	ICC	CoV	ICC	CoV
0.01	0.90	9.8	0.84	11.9	0.98	5.5
0.02	0.96	6.2	0.89	9.5	0.98	5.8
0.03	0.97	5.1	0.95	6.5	0.99	3.7
0.04	0.97	5.3	0.94	6.8	0.99	4.1
0.05	0.98	3.9	0.96	5.4	0.99	3.5
0.06	0.98	3.7	0.95	5.9	0.99	3.4
0.07	0.98	3.9	0.96	5.1	0.99	3.4
0.08	0.99	3.4	0.96	5.2	0.99	3.1
0.09	0.99	3.4	0.97	4.9	0.99	3.1

* ICC as well as CoV values are provided as a function of Δz. The center point is located at $z_0 = 0.8$, $\theta_0 = 210^\circ$ and $\Delta\theta = 100^\circ$. The intra and inter-reader precision was obtained from subset of the FNIH data used for the main analysis. The repositioning reproducibility data were measured using a different set of OAI subjects where patient repositioning was done.

appropriate for existing and future large studies of knee OA that use high resolution MRI.

The LACS method has good responsiveness with a 24 month SRM as high as −0.52. Previous work in a 24 month follow-up study using this segmentation method on the same femoral region showed similar SRMs[10]. A significantly more responsive result was found (SRMs up to −1.21)[19] when using a 3D image registration method applied to a focal area of cartilage centered at a unique indexed point chosen by a reader as the area of greatest thinning on the follow-up image. For future studies, it will be possible to use the coordinate system from the LACS method on indexed regions by calculating the location indicated by the reader in z and θ for both time points. By providing a unique indexed location for each subject, this method could potentially offer a measurement that is as fast, but more responsive than LACS alone.

We assessed cartilage volume in the central medial weight-bearing region of the femur and showed significant results for the primary case–control analysis of subjects with radiographic and pain progression combined (Group 1) vs subjects with radiographic progression only, subjects with pain progression only, and subjects with neither radiographic nor pain progression (Group 2, 3, and 4) ($\Delta z = 0.09$: OR = 1.70; 95% CI = 1.40–2.07). There was an even stronger association between cartilage volume measurements and radiographic progression alone (Group 2 vs Group 4). For the analysis of subjects with pain progression only (Group 3) vs subjects with neither radiographic nor pain progression (Group 4) the ORs are lower than in the other comparisons and do not reach significance. This is in concordance with most other studies that report a weak association between structural measures of OA and pain[20].

The OR data are consistent with the results from an independent study[13], which used the same FNIH scans but a different segmentation technique that measured cartilage thickness. In this independent study, the authors found the highest ORs in a central weight-bearing portion of the femur that is in a similar location as the LACS region. For this reason, we would expect lower ORs and responsiveness for other regions of the femur segmented with LACS. Reader times are not quoted for the independent study[13].

A key finding of our study is that different sized regions of cartilage segmentation yielded mostly similar ORs and SRMs apart from very small areas (smaller than Δz of 0.03). As discussed previously, this implies that substantially faster reader times will be possible for future studies. Currently, the largest area ($\Delta z = 0.09$, $\Delta\theta = 100^\circ$) of cartilage segmentation required approximately 20 min for segmentation. Reduction in Δz corresponds directly (in a linear fashion) into fewer total slices presented to the reader, and therefore, to proportionally reduced time, since the reader reviews one slice at a time. With this in mind, using a region defined by $\Delta z = 0.03$ would require less than 7 min of total reader time, while providing a method that is nearly as responsive and discriminatory as with the larger region. Such an approach would facilitate studies involving tens of thousands of scans such as the OAI. Various methods are reported in the literature with different levels of automation. In general, increased automation leads to a reduction in reader time. The LACS technique is substantially automated and maintains the same level of performance as manual segmentation[13].

As expected, intra-reader was superior to inter-reader precision and, as with the ORs and responsiveness, performance was relatively unchanged until low values of Δz. Repositioning reproducibility for the 15 scan pairs showed a similar trend. An independent study, using these images to measure the full femur cartilage volume, found a repositioning reducibility of CoV = 2.0, which can be compared to CoV = 3.1 for the LACS method. Since the average volume for the LACS method is much smaller, the CoV may not be an ideal metric for comparing these two quantities due to the denominator effect.

For this study, we have selected a region in the medial compartment central weight-bearing portion of the femur, which may be less relevant for advanced OA since it is the most frequent site of cartilage loss. For future studies, different areas in the femur may be of greater interest. The method can easily be applied to any individual region of the femur as well as multiple sites covering several distinct locations. Given that that performance is maintained for even very small regions of cartilage, the reader time for each of the multiple measurements could potentially be very low,

and the total time correspondingly moderate. Assessing several smaller regions would be faster than the full femur since a substantial amount of the cartilage plate would not have to be segmented.

Limitations of our approach are that at this stage we are only measured a focused area of cartilage in the medial femur and not in the patella or the tibia. The specific method we used for this study may not capture cartilage loss associated with other areas of the femur or the lateral compartment. However, in its current form, the LACS method can easily be extended to assess any region of the femur. In the future, measurements of cartilage volume in the tibia and patella will be possible with an analogous coordinate system. We currently only provide our measurements in the 3D DESS sequence. The method we used in this study does not provide measurements of cartilage volume of the whole cartilage plate even in the medial femur. But our method delivers high SRMs and significant ORs, indicating that this focused approach provides not only responsive but also fast measurements applicable for studies with thousands of knee MRIs to assess. The OA Biomarkers Consortium study cohort represents a preselected study cohort that does not necessarily reflect the general population, but it might be similar to other study cohorts for which this method is intended.

Conclusions

We further validated a semi-automated, quantitative tool to measure cartilage volume in the central medial weight-bearing region of the femur and demonstrated clinical validity in a case–control sample of knee OA progression. Furthermore we showed that equivalent performance is possible with much shorter reader times.

Author contributions

Study conception and design – *Schaefer, Lynch, Duryea*.

Acquisition of data – Duryea, Schaefer, Jamieson, Donnell, Sury, Yin.

Analysis and interpretation of data – *Duryea, Schaefer, Smith, Lynch, Nevitt*.

Drafting the article or revising it critically for important intellectual content – *Duryea, Schaefer, Smith, Jamieson*, Donnell, *Sury, Yin, Lynch, Nevitt*.

Final approval of the version of the article to be published – *Duryea, Schaefer, Smith*, *Jamieson*, Sury, *Yin, Lynch, Nevitt*.

L. Schaefer (lenafranziskaschaefer@yahoo.com) takes responsibility for the integrity of the work as a whole.

Conflict of interest

The authors declare that they have no conflicting interests.

Role of funding source

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, a branch of the U.S. Department of Health and Human Services, 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. This study was supported by the NIH/NIAMS (R01AR056664).

Acknowledgements

We acknowledge Quinley Miao for her invaluable assistance with the image analysis setup. This study was supported by the NIH/NIAMS (R01AR056664). Data provided from the FNIH OA Biomarkers Consortium Project are made possible through grants and direct or in-kind contributions by: AbbVie; Amgen; Arthritis Foundation; Artialis; Bioiberica; BioVendor; DePuy; Flexion Therapeutics; GSK; IBEX; IDS; Merck Serono; Quidel; Rottapharm | Madaus; Sanofi; Stryker; the Pivotal OAI MRI Analyses (POMA) study, NIH HHSN2682010000 21C; and the Osteoarthritis Research Society International.

References

1. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis. Nat Rev Rheumatol 2014;10:437–41.
2. Hayashi D, Roemer FW, Guermazi A. Imaging for osteoarthritis. Ann Phys Rehabil Med 2016;59:161–9.
3. Lester G. Clinical research in OA–the NIH osteoarthritis initiative. J Musculoskelet Neuronal Interact 2008;8:313–4.
4. 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.
5. Stefanik JJ, Guermazi A, Roemer FW, Peat G, Niu J, Segal NA, *et al*. Changes in patellofemoral and tibiofemoral joint cartilage damage and bone marrow lesions over 7 years: the Multicenter Osteoarthritis Study. Osteoarthritis Cartilage 2016;24:1160–6.
6. Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). Ann Rheum Dis 2008;67:206–11.
7. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, *et al*. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage 2011;19:990–1002.
8. Eckstein F, Guermazi A, Gold G, Duryea J, Hellio Le Graverand MP, Wirth W, *et al*. Imaging of cartilage and bone: promises and pitfalls in clinical trials of osteoarthritis. Osteoarthritis Cartilage 2014;22:1516–32.
9. Wildi LM, Martel-Pelletier J, Abram F, Moser T, Raynauld JP, Pelletier JP. Assessment of cartilage changes over time in knee osteoarthritis disease-modifying osteoarthritis drug trials using semiquantitative and quantitative methods: pros and cons. Arthritis Care Res (Hoboken) 2013;65:686–94.
10. Duryea J, Iranpour-Boroujeni T, Collins JE, Vanwynngaarden 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;66:1560–5.
11. Duryea J, Cheng C, Schaefer LF, Smith S, Madore B. Integration of accelerated MRI and post-processing software: a promising method for studies of knee osteoarthritis. Osteoarthritis Cartilage 2016;24:1905–9.
12. Hunter DJ, Nevitt M, Losina E, Kraus V. Biomarkers for osteoarthritis: current position and steps towards further validation. Best Pract Res Clin Rheumatol 2014;28:61–71.
13. Eckstein F, Collins JE, Nevitt MC, Lynch JA, Kraus VB, Katz JN, *et al*. Brief report: cartilage thickness change as an imaging biomarker of knee osteoarthritis progression: data from the

Foundation for the National Institutes of Health osteoarthritis biomarkers consortium. Arthritis Rheumatol 2015;67:3184–9.
14. Hunter D, Nevitt M, Lynch J, Kraus VB, Katz JN, Collins JE, *et al.* Longitudinal validation of periarticular bone area and 3D shape as biomarkers for knee OA progression? Data from the FNIH OA Biomarkers Consortium. Ann Rheum Dis 2016;75:1607–14.
15. Peterfy C, Li J, Zaim S, Duryea J, Lynch JA, Miaux Y, *et al.* Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-space width in the knee: test-retest reproducibility. Skelet Radiol 2003;3:128–32.
16. Collins JE, Losina E, Nevitt MC, Roemer FW, Guermazi A, Lynch JA, *et al.* Semi-quantitative imaging biomarkers of knee osteoarthritis progression: data from the FNIH OA biomarkers consortium. Arthritis Rheumatol 2016;68:2422–31.
17. Duryea J, Neumann G, Brem MH, Koh W, Noorbakhsh F, Jackson RD, *et al.* Novel fast semi-automated software to segment cartilage for knee MR acquisitions. Osteoarthritis Cartilage 2006;15:487–92.
18. Eckstein F, Hudelmaier M, Wirth W, Kiefer B, Jackson R, Yu J, *et al.* Double echo steady state magnetic resonance imaging of knee articular cartilage at 3 Tesla: a pilot study for the Osteoarthritis Initiative. Ann Rheum Dis 2006;65:433–41.
19. Iranpour-Boroujeni T, Watanabe A, Bashtar R, Yoshioka H, Duryea J. Quantification of cartilage loss in local regions of knee joints using semi-automated segmentation software: analysis of longitudinal data from the Osteoarthritis Initiative (OAI). Osteoarthritis Cartilage 2011;19:309–14.
20. Hunter DJ, Guermazi A, Roemer F, Zhang Y, Neogi T. Structural correlates of pain in joints with osteoarthritis. Osteoarthritis Cartilage 2013;21:1170–8.