

Osteoarthritis and Cartilage



Can cartilage loss be detected in knee osteoarthritis (OA) patients with 3–6 months' observation using advanced image analysis of 3 T MRI?

D.J. Hunter ^{†*}, M.A. Bowes [‡], C.B. Eaton ^{§||}, A.P. Holmes [¶], H. Mann [¶], C.K. Kwok ^{#††}, R.A. Maciewicz [¶], J. Samuels ^{‡‡}, J.C. Waterton [¶]

[†] Division of Research, New England Baptist Hospital, 125 Parker Hill Ave, Boston, MA 02120, USA

[‡] Imorphics, Kilburn House, Manchester Science Park, Manchester M15 6SE, UK

[§] Center for Primary Care and Prevention, Memorial Hospital of Rhode Island, 111 Brewster St, Pawtucket, RI 02860, USA

^{||} Department of Family Medicine and Community Health (Epidemiology), Alpert Medical School of Brown University, Pawtucket, RI 02860, USA

[¶] AstraZeneca, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK

[#] Division of Rheumatology and Clinical Immunology, University of Pittsburgh, 3500 Terrace St, Pittsburgh, PA 15261, USA

^{††} Pittsburgh VA Healthcare System, Pittsburgh, PA 15206, USA

^{‡‡} Seligman Center for Advanced Therapeutics, NYU Hospital for Joint Diseases, 301 East 17th St, New York, NY 10003, USA

ARTICLE INFO

Article history:

Received 29 June 2009

Accepted 1 February 2010

Keywords:

Osteoarthritis

Magnetic resonance imaging

Cartilage thickness

SUMMARY

Purpose: Prior investigations of magnetic resonance imaging (MRI) biomarkers of cartilage loss in knee osteoarthritis (OA) suggest that trials of interventions which affect this biomarker with adequate statistical power would require large clinical studies of 1–2 years duration. We hypothesized that smaller, shorter duration, “Proof of Concept” (PoC) studies might be achievable by: (1) selecting a population at high risk of rapid medial tibio-femoral (TF) progression, in conjunction with; (2) high-field MRI (3 T), and; (3) using advanced image analysis. The primary outcome was the cartilage thickness in the central medial femur.

Methods: Multi-centre, non-randomized, observational cohort study at four sites in the US. Eligible participants were females with knee pain, a body mass index (BMI) ≥ 25 kg/m², symptomatic radiographic evidence of medial TF OA, and varus mal-alignment. The 29 participants had a mean age of 62 years, mean BMI of 36 kg/m², with eight index knees graded as Kellgren–Lawrence (K&L) = 2 and 21 as K&L = 3. Eligible participants had four MRI scans of one knee: two MRIs (1 week apart) were acquired as a baseline with follow-up MRI at 3 and 6 months. A trained operator, blind to time-point but not subject, manually segmented the cartilage from the Dual Echo Steady State water excitation MR images. Anatomically corresponding regions of interest were identified on each image by using a three-dimensional statistical shape model of the endosteal bone surface, and the cartilage thickness (with areas denuded of cartilage included as having zero thickness – ThCtAB) within each region was calculated. The percentage change from baseline at 3 and 6 months was assessed using a log-scale analysis of variance (ANOVA) model including baseline as a covariate. The primary outcome was the change in cartilage thickness within the aspect of central medial femoral condyle exposed within the meniscal window (w) during articulation, neglecting cartilage edges [nuclear (n)] (nwcMF·ThCtAB), with changes in other regions considered as secondary endpoints.

Results: Anatomical mal-alignment ranged from -1.9° to 6.3° , with mean 0.9° . With one exception, no changes in ThCtAB were detected at the 5% level for any of the regions of interest on the TF joint at 3 or 6 months of follow-up. The change in the primary variable (nwcMF·ThCtAB) from (mean) baseline at 3 months from the log-scale ANOVA model was -2.1% [95% confidence interval (CI) $(-4.4\%, +0.2\%)$]. The change over 6 months was 0.0% [95% CI $(-2.7\%, +2.8\%)$]. The 95% CI for the change from baseline did not include zero for the cartilage thickness within the meniscal window of the lateral tibia (wLT·ThCtAB) at 6 month follow-up (-1.5% , 95% CI $[-2.9, -0.2]$), but was not significant at the 5% level after correction for multiple comparisons.

* Address correspondence and reprint requests to: David John Hunter, Division of Research, New England Baptist Hospital, 125 Parker Hill Ave, Boston, MA 02120, USA. Tel: 1-617-754-6655; Fax: 1-617-754-6888.

E-mail address: djhunter@caregroup.harvard.edu (D.J. Hunter).

Conclusions: The small inconsistent compartment changes, and the relatively high variabilities in cartilage thickness changes seen over time in this study, provide no additional confidence for a 3- or 6-month PoC study using a patient population selected on the basis of risk for rapid progression with the MRI acquisition and analyses employed.

© 2010 Osteoarthritis Research Society International. Published by Elsevier Ltd. All rights reserved.

Introduction

One proposed osteoarthritis (OA) treatment goal is preservation of the underlying joint structure. In an effort to shorten development timelines, and minimize patient exposure to investigational therapies, clinical trial brevity is paramount. As OA is typically a very slowly progressive condition, one can optimize trial efficiency by finding more responsive endpoint/s and/or stratifying the study sample to further enhance efficiency.

At the time this study was conceived in 2006, consensus in the literature suggested that changes in magnetic resonance imaging (MRI) measures of cartilage morphology ≥ 1 year were of the order of 4–6% in selected patient groups^{1–4}. Recent studies have demonstrated superior responsiveness in the central medial femur than the medial tibia^{5,6}. However these data sets did not evaluate intervals shorter than 6 months, and thus it was not clear if one could conduct “Proof of Concept” (PoC) studies over shorter time intervals. The study sponsor had recently conducted a 6-month follow-up in elderly, obese, female participants with radiographic evidence of OA [Kellgren–Lawrence (K&L) grade 2 or 3] and current knee pain⁷. The compartmental cartilage volumes analyzed in this 6-month study showed no significant changes after 6 months. However, analysis of cartilage thickness changes within smaller focal regions identified using three-dimensional (3D) statistical shape modelling giving an Anatomically Corresponded Regional Analysis of Cartilage (ACRAC) detected decrease in cartilage thickness within the trimmed central medial femoral region of 8.2%, with 95% confidence interval (CI) (–15.6%, –0.2%)⁷.

In addition to image analysis techniques, several methods of stratifying populations to enrich for persons at greater risk of medial tibio-femoral (TF) compartment progression have been proposed. These include increased body mass index (BMI)⁸, an increased level of type II collagen C-terminal degradation products detected in the urine⁹, the presence of varus mal-alignment at the TF joint^{10–14}, the presence on MRI of sub-chondral bone marrow lesions¹⁵ or meniscal abnormalities¹⁶.

The primary objective of this study was to investigate the potential of central medial femoral cartilage thickness, measured by optimal MRI sequences with advanced image analysis, as a biomarker endpoint for future 3-month disease modification PoC studies. The secondary objective was to investigate the potential of changes in cartilage thickness in any region over 3 or 6 months measured by MRI as biomarker endpoints for an OA PoC study.

Materials and methods

Study design

This was a multi-centre, non-randomized, observational cohort (Level II) study to assess the utility of MRI in detecting morphological changes in knee OA over 3 and 6 months. One hundred and forty-one participants were enrolled across four sites in the US. Following screening assessments and confirmation of eligibility against all but the X-ray inclusion criteria at Visit 1, participants were scheduled to attend Visit 2a and arrangements made for a fixed-flexion posterior-anterior (PA) view knee X-ray to be obtained and centrally assessed. At Visit 2a, following confirmation of X-ray eligibility criteria from the

central review, participants were accepted for follow-up and the target knee joint selected. The initial baseline MRI scan of the target knee was then arranged to take place within 3 days of Visit 2a. Subsequently, participants were to be followed-up 1 week \pm 2 days (Visit 2b), three calendar months \pm 7 days (Visit 3) and six calendar months \pm 7 days (Visit 4) later (see Table I).

After each scheduled visit, MRI scanning of the target knee was to be performed within 3 days. The Visit 2b, three and four MRI scans were stipulated to take place at the same time of day (\pm 2 h) as the first MRI scan at Visit 2a. Participants were to remain rested for a minimum of 30 min prior to each MRI scan (i.e., seated in the MRI clinic waiting room).

Eligibility criteria

We targeted recruitment at participants with medial TF OA and varus mal-alignment of the knee. The inclusion criteria assessed at Visit 1 included:

1. Female participants aged 50 years and above.
2. BMI of 25 kg/m² or above.
3. Radiographic evidence (within the last 3 months) of TF OA defined as fixed-flexion PA radiographs of the target knee(s) showing definite medial TF osteophytes (OARSI grades 1–3) and joint space narrowing (JSN) (OARSI grades 1–2, medial JSN \geq lateral JSN), excluding participants with severe JSN (OARSI grade 3 or bone on bone), as assessed by central reading of the X-ray¹⁷.
4. Knee varus mal-alignment ($\geq -2^\circ$ according to anatomical alignment), determined by central reading of the fixed-flexion radiograph (PA view)^{18,19}.
5. Symptomatic disease as defined by the presence of “pain, aching or stiffness in or around the knee on most days” for at least 1 month during the past 12 months.
6. Current knee pain, defined as knee pain, in either knee, in the one week before Visit 1, for which the patient gives a score of at least 3 on a 0–10 scale for global knee pain (where 0 = no pain, and 10 = worst pain imaginable).

Exclusion criteria assessed at Visit 1 included:

1. Diagnosis of secondary OA due to inflammatory joint disease, Paget’s disease of the bone, major dysplasias or congenital abnormality, or current diagnosis of another form of arthritis in addition to OA.
2. Insufficient knee flexion for standardized positioning in X-ray.
3. Any planned surgery within the next 6 months.
4. Intra-articular injections of hyaluronic acid or corticosteroids to the knee or other invasive orthopaedic surgery in the last 6 months, or planned within the next 6 months.
5. Major surgery or significant trauma within the last 3 months.

Table I

Study timeline and interval between visits

Visit 1	Visit 2a	Visit 2b	Visit 3	Visit 4
Screening	Baseline	1 week	3 months	6 months

6. A history of claustrophobia or any other contra-indications to the practical aspects of MRI scanning.
7. Risk (in the investigator's opinion) of transmitting HIV or Hepatitis B.
8. Pregnancy.
9. Participation in another clinical study involving an investigational compound, the last follow-up visit of which was within 90 days of Visit 1 in this study.
10. Use of ambulatory aids, other than a single cane, for more than 50% of the time in ambulation.
11. Current use of Disease Modifying OA Drugs (DMOADs) (e.g., doxycycline, diacerin, glucosamine, chondroitin) unless on dose which had been stable for at least 3 months prior to Visit 2, and which would remain stable during the study.

Plain radiographs

Participants were required to have a fixed-flexion knee X-ray (PA view) to confirm eligibility for this study. PA views were obtained using a SynaFlexer™ frame (Synarc, Inc., San Francisco, CA) to position the subject's feet reproducibly²⁰.

X-ray images were assessed by a musculoskeletal radiologist to evaluate for individual radiographic features of JSN, and osteophyte grading according to the OARSI atlas¹⁷, K&L grade²¹, and anatomical alignment. The anatomic axis was determined as the angle formed by the intersection of two lines originating from points bisecting the femur and tibia and converging at the centre of the tibial spine tips (inter-condylar eminence), using three points to specify the lines as previously described¹⁸.

MRI acquisition

MRI was performed on the index knee using Siemens 3 T Trio systems using the same imaging sequences as used in the OA Initiative (OAI)²². MR imaging centres conducted pre-study training that included safety pre-screening, subject identification coding, loading of protocols into the MRI system, scanning, data transfer and site archival procedures. Uniformity and Linearity (UAL) and knee phantoms supplied by VirtualScopics (Rochester, NY) were scanned and visually assessed for distortion, warping and double shadowing along with protocol compliance. Further quality control continued for the duration of the study.

MR image analysis

The MR images were transferred to Imorphics for analysis using the Anatomically Corresponded Regional Analysis of Cartilage (ACRAC) technique²³. Cartilage was segmented from the Dual Echo Steady State water excitation (DESSwe) MR images by a single trained operator (BW, acknowledged below) who was blind to visit but not to patient. The segmenter had previously segmented two full scale trials of OAI data, and had passed the Imorphics training protocol, which requires segmenters to be able to repeatedly segment the femoral, medial tibial, lateral tibial and patella cartilage compartments an intra-observer coefficient of variation (CoV) of less than 3% using a set of paired images blinded at random. The cartilage outline was identified on each slice by manual tracing using Imorphics EndPoint software, following the Imorphics segmentation protocol. For each patient, the MR images were segmented in pairs, using the first image segmented as a reference for further segmentations. The first image to be segmented was allocated at random. The remaining images were segmented in turn with the first segmentation from the patient for reference. All of the hyaline cartilage was identified

in each slice of each image for the femoral, medial tibial, lateral tibial and patellar cartilage.

Cartilage segmentations were reviewed by an expert segmenter with 7 years experience of cartilage segmentation (MB) for consistency. The reviewer was also blinded to visit but not to patient. In the event that the expert segmenter was uncertain of a reading, a musculoskeletal radiologist (CH) was consulted, and the segmentation agreed.

3D cartilage surfaces were constructed from the planar contours using shape interpolation between adjacent slices, utilizing a 3D signed-distance function using information from each planar contour. The 3D surfaces were reviewed visually against the original image, and using 3D VRML visualization.

A statistical model was built from the DESSwe images of both knees from the 160 participants initially released by the OAI as group 0.B.1. The models were built using the piecewise affine registration method²⁴ which takes the whole of the MRI image around the knee for each member of the training set, and generates a statistical model for the whole volume. During model construction, each member of the training set is populated with a dense network of control points throughout the volume.

This model was used to generate a mean image of the 160 subjects, which was manually segmented by a musculoskeletal radiologist (CH) to identify the bone surface of the femur, tibia and patella. The articular cartilage and the menisci were also segmented.

This segmentation was then used to create 3D surfaces for the bone, meniscus and cartilage, which can then be projected to all examples in the model, using the control points. In the femur over 30,000 points are fitted to the bone surface – previous studies have shown that when these points are propagated out to the examples in the training set, they have mean positional errors of less than 1 mm²⁵. This, therefore provides a dense, anatomical corresponded set of points, which can be used to take measurements and identify regions for any image to which the bone model is fitted.

The tAB regions for the femur, tibia and patella were then identified, defined as the bone which is covered by cartilage within the mean image. Similarly, the meniscal window on the tibia (w) can be identified as the region on the bone surface which falls within the menisci in the image.

On the tibia, the meniscal window is not subdivided, but on the femur, we further divide the meniscal window into a central section (c) and a posterior section (p). A line is drawn along the bone which is directly beneath the posterior edge of the meniscus; this line is extended smoothly both medially and laterally to the edge of the tAB region. The central section is defined as the area of bone anterior to this line, and the posterior region is the windowed region behind this (Fig. 2).

To consistently remove the cartilage edges, which are most prone to noise and partial volume, a trimming boundary was

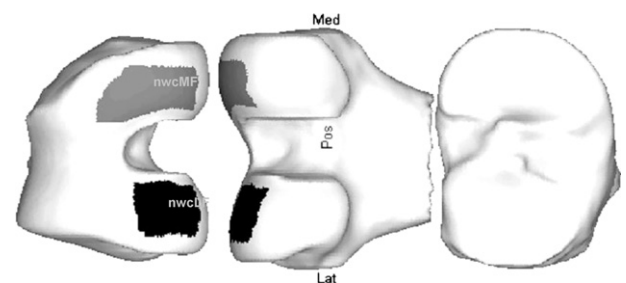


Fig. 1. Regions of interest: nuclear (trimmed) central meniscal window regions on the medial and LF (nwcMF & nwcLF).

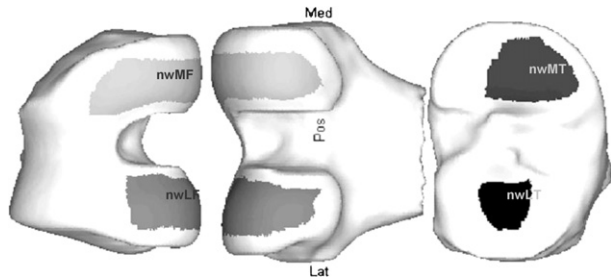


Fig. 2. Regions of interest: nuclear (trimmed) meniscal window regions (nwMF, nwLF, nwMT, nwLT).

defined within the model. A series of regions were generated on the mean bone model using an algorithm that progressively eroded the tAB areas evenly from their edges. This was achieved by selecting a new series of correspondence points with approximately the same shape as the original region. Because the trimming boundary is formed from a line drawn through correspondence points on the model, the trimming boundary is propagated automatically to each image when the model was fitted. The trimming boundary that was chosen for the femoral, medial and lateral tibial and patellar cartilage plate reduced each of them by around 35%.

Each image in the study was then fitted with the bone reference surface from the model, thereby also identifying the tAB regions, and the various anatomical regions described above.

The primary outcome variable was the overall mean cartilage thickness (with areas denuded of cartilage included as having zero thickness – ThCtAB²⁶) within the aspect of the central (c) medial femur (MF) condyle which is exposed within the meniscal window (w) during articulation with the cartilage edges “trimmed” leaving the nuclear (n) region as described above, with nomenclature nwcMF.

The secondary variables were the overall mean cartilage thickness (with denuded areas included as having zero thickness – ThCtAB) across the following regions (see Figs. 1 and 2 for examples)²⁶:

- On the medial femur (seven regions): medial femur (MF), central medial femur (cMF), medial femur meniscal window (wMF), central medial femur meniscal window (wcMF), nuclear (trimmed) medial femur (nMF), nuclear (trimmed) central medial femur (ncMF), nuclear (trimmed) medial femoral meniscal window (nwMF).
- On the medial tibia (four regions): medial tibia (MT), medial tibial meniscal window (wMT), nuclear medial tibia (nMT), nuclear medial tibial meniscal window (nwMT).
- Across the medial TF joint (four regions): central medial TF (cMTF), central medial TF meniscal window (wcMTF), nuclear central medial TF (ncMTF), nuclear central medial TF meniscal window (nwcMTF).
- On the lateral femur (LF) (eight regions): LF, central LF (cLF), lateral femur meniscal window (wLF), central lateral femur meniscal window (wcLF), nuclear (trimmed) LF (nLF), nuclear (trimmed) central LF (ncLF), nuclear (trimmed) lateral femur meniscal window (nwLF), nuclear (trimmed) central lateral femur meniscal window (nwcLF).
- On the lateral tibia (LT) (four regions): LT, lateral tibial meniscal window (wLT), nuclear LT (nLT), nuclear lateral tibial meniscal window (nwLT).
- Across the lateral TF (LTF) joint (four regions): central LTF (cLTF), central LTF meniscal window (wcLTF), nuclear central LTF (ncLTF), nuclear central LTF meniscal window (nwcLTF).

Baseline MRI cartilage morphology measures

Two baseline MRI assessments were carried out a week apart (± 2 days) at Visits 2a and 2b to assess the reproducibility of the technique. The mean of the values from these two baseline assessments or the single value (where only one value was available), referred to as “mean baseline” was used as the baseline value from which change at 3 and 6 months was assessed. The mean of the two baseline values was calculated on the logarithmic scale for analyzing proportional change and calculating changes as a percentage of baseline, and is therefore the geometric mean on the natural scale. The averaging of the two baseline scans can reduce the variability, potentially yielding more accurate measurements of change and thereby improving the power of the analysis in comparison to using the value from only one of the Visit two MRI assessments as baseline value.

Statistical analysis methods

The analysis set is the set of participants who completed Visit 2a, had at least one follow-up visit (2b, 3 or 4), and were not subsequently withdrawn due to protocol violations.

The statistical analysis for the primary and secondary objective was performed using an analysis of variance (ANOVA) model. Since it is the rate of change in the primary variable at 3 months from baseline, which is of primary interest, the ANOVA response variable was the log transformed ratio of the 3 months to the baseline measurement. The model included a fixed factor for visit and a baseline covariate for the (log transformed) cartilage thickness value. The mean change and 95% CI were back-transformed for presentation in terms of the percentage change from baseline, along with the corresponding *P*-value.

Assumptions of normality were explored graphically through normal probability plots for the residuals from fitting this model. If assessment of the model residuals for the primary variable suggested that the model is a reasonable fit to the data, then the same model was to be used to analyze all the MRI cartilage morphology variables. If not, a non-parametric technique or the actual change from baseline was to be considered to validate the results of the main analysis.

The level of agreement between the baseline values for the primary variable was assessed across the patient group using the statistical techniques of Bland–Altman for method comparison studies²⁷, i.e., by plotting the difference between the two baseline measurements against the mean of the measurements and including a reference range of 95% limits of agreement given by $d \pm 1.96s$ where d is the mean difference and s is the standard deviation of the difference between the two baseline measurements. The baseline reproducibility of each variable was obtained by fitting an ANOVA model to the log transformed variable including subject as a random effect. The inter-subject components of variation on the log-scale were obtained from the ANOVA models and are presented back-transformed as CoV.

Results

One hundred and forty-one participants were enrolled across four sites in the US. Omitting participants who were withdrawn due to ineligibility, 30 participants completed Visit 2a and continued in the study. Of these, one participant was lost to follow-up after Visit 2a and had no further assessments. The remaining 29 had at least one follow-up visit.

Table II shows the baseline characteristics of the study sample. The mean age was 62.2 years, with range 50–80 years. The mean BMI was 35.9, with range 30.8–50.9. Two participants were black,

Table II
Baseline characteristics of study sample

Gender [female N (%)]	29 (100)	
Age (mean, SD), years	62.1 (8.3)	
BMI (mean, SD), kg/m ²	35.9 (4.8)	
Index knee [left N (%)]	14 (48)	
WOMAC pain (mean, SD) (0–20 scale)	7.3 (2.6)	
K&L grade of index knee, no. (%)	Grade	Number of knees
	2	8 (28)
	3	21 (72)
Anatomic axis (mean, SD),°	0.91 (2.0)	

the remaining 27 were white. Eight subjects had a K&L grade of 2, the remaining 21 had a grade of 3. Anatomical alignment ranged from -1.92° to 6.27° , with mean 0.91° , where varus mal-alignment is measured in the positive direction.

Twenty-seven of the 29 participants in the overall analysis set had overall mean cartilage thickness (with areas denuded of cartilage included as having zero thickness) (ThCtAB) within the aspect of central medial femoral condyle exposed within the meniscal window during articulation with cartilage edges “trimmed” (nwcMF) measured from MRI scans of sufficient quality at least one baseline visit (Visit 2a or Visit 2b) and at the 3-month follow-up visit (Visit 3) and were included in the analysis. The final analysis sets for the MRI objectives consisted of 28 subjects for objectives not involving the 3-month time-point, with one less (27) subjects for the assessment of 3-month change and linearity.

The change in the primary variable (nwcMF·ThCtAB) from (mean) baseline at 3 months from the log-scale ANOVA model was -2.1% , 95% CI (-4.4% , $+0.2\%$). The change over 6 months was 0.0% (95% CI $[-2.7\%$, $+2.8\%]$).

Changes in the primary and secondary MRI cartilage thickness variables over 3 and 6 months from (mean) baseline calculated from the ANOVA model are tabulated in Table III. Results for the nuclear (trimmed) central meniscal windowed regions on the medial and LT and femur (nwcMF, nwMT, nwcLF, nwLT) are plotted in Fig. 3. The 95% CI for the change from baseline did not include zero for the cartilage thickness within the meniscal window of the LT over 6 months [-1.5% , 95% CI (-2.9% , -0.2%)]. With that exception, no significant changes in cartilage thickness were detected at the 5% level for any of the regions of interest on the TF joint over 3 or 6 months in this study.

Twenty-eight of the 29 participants in the overall analysis set had the primary MRI cartilage morphology variable measured from MRI scans of sufficient quality at both baseline visits (Visit 2a and Visit 2b) and were included in the analysis of baseline reproducibility.

The Bland–Altman plots for the primary variable nwcMF·ThCtAB at Visits 2a and 2b show limits of agreement between the two baseline visits of approximately ± 0.13 mm (Fig. 4). However, the mean cartilage thickness over this region for all 28 subjects is only 1.5 mm at the baseline visits, and calculated on the log-scale the limits of agreement as percentages are (-15% , 17%) (Fig. 5).

The repeatability of the measurement of cartilage thickness was typically around 2% measured as CoV (RMS). This compares well with other studies²⁸ and indicates that the measurement noise in this study is at least as small as published work (Table IV). The repeatability of the bone area tAB is provided in the second column. This shows the repeatability of the bone search, and its ability to automatically identify a propagated region. The repeatability of the bone search is considerably better than that of the cartilage thickness. It is notable that the repeatability of the primary region nwcMF has the poorest repeatability (though still comparable with other authors). This may be caused by the presence of denuded cartilage in this region, which is usually the source of increased noise in cartilage measurement.

Table III
Changes from mean baseline at 3 and 6 months in MRI cartilage thickness variables

Region	Over 3 months			Over 6 months		
	N	% Change	95% CI	N	% Change	95% CI
nwcMF	27	-2.1	(-4.4, 0.2)	28	0.0	(-2.7, 2.8)
nwMF	27	-0.9	(-2.1, 0.2)	28	1.1	(-0.4, 2.6)
ncMF	27	-1.3	(-2.6, 0.1)	28	0.6	(-1.3, 2.5)
nMF	27	-0.8	(-1.7, 0.1)	28	1.3	(-0.1, 2.6)
wcMF	27	-1.6	(-3.7, 0.5)	28	0.4	(-2.0, 2.8)
wMF	27	-1.0	(-2.5, 0.5)	28	1.2	(-0.2, 2.7)
cMF	27	-1.1	(-2.6, 0.4)	28	0.6	(-1.1, 2.3)
MF	27	-1.1	(-2.4, 0.1)	28	1.0	(-0.6, 2.5)
nwMT	27	0.7	(-2.3, 3.8)	28	-1.2	(-3.3, 1.0)
nMT	27	0.4	(-2.1, 2.9)	28	-0.5	(-2.9, 2.0)
wMT	27	0.8	(-2.0, 3.6)	28	-1.0	(-2.9, 0.9)
MT	27	0.3	(-2.0, 2.8)	28	-1.4	(-3.5, 0.8)
nwcMTF	27	-0.5	(-2.2, 1.3)	28	-0.4	(-2.3, 1.6)
ncMTF	27	-0.4	(-1.7, 1.0)	28	0.2	(-1.7, 2.0)
wcMTF	27	-0.4	(-2.0, 1.3)	28	-0.1	(-1.8, 1.6)
cMTF	27	-0.4	(-1.8, 1.0)	28	-0.3	(-1.9, 1.3)
nwcLF	27	-0.1	(-1.2, 0.9)	28	0.3	(-0.8, 1.5)
nwLF	27	0.7	(-0.5, 2.0)	28	0.3	(-1.0, 1.5)
ncLF	27	0.1	(-0.8, 1.1)	28	-0.1	(-1.3, 1.1)
nLF	27	0.8	(-0.4, 2.0)	28	-0.1	(-1.5, 1.3)
wcLF	27	0.0	(-1.1, 1.2)	28	0.4	(-0.7, 1.5)
wLF	27	0.5	(-0.8, 1.9)	28	0.2	(-1.0, 1.5)
cLF	27	0.2	(-1.1, 1.6)	28	-0.3	(-1.5, 0.9)
LF	27	0.7	(-1.0, 2.4)	28	-0.6	(-1.9, 0.9)
nwLT	27	-0.1	(-2.0, 1.8)	28	-1.2	(-2.4, 0.1)
nLT	27	0.7	(-0.4, 1.7)	28	-0.6	(-1.7, 0.5)
wLT	27	-0.2	(-2.3, 2.0)	28	-1.5	(-2.9, -0.2)
LT	27	1.1	(-0.1, 2.3)	28	-0.9	(-2.3, 0.6)
nwcLTF	27	-0.2	(-1.2, 0.9)	28	-0.1	(-1.0, 0.7)
ncLTF	27	0.3	(-0.5, 1.2)	28	-0.3	(-1.1, 0.5)
wcLTF	27	-0.1	(-1.2, 1.1)	28	-0.2	(-1.1, 0.6)
cLTF	27	0.6	(-0.5, 1.8)	28	-0.5	(-1.4, 0.4)

The mean and 95% CI estimates are from fitting ANOVA models to the log transformed data.

The ANOVA models include the baseline measure as a covariate.

The baseline is the geometric mean of the value at Visits 2a and 2b.

CI s are not corrected for multiple comparisons.

Legend for regions as detailed in Materials and methods section.

Discussion

With one exception (lateral tibial meniscal window), no significant changes in cartilage thickness were detected at the 5% level for any of the regions of interest on the TF joint over 3 or 6 months in this study. However, the changes in 32 primary and secondary

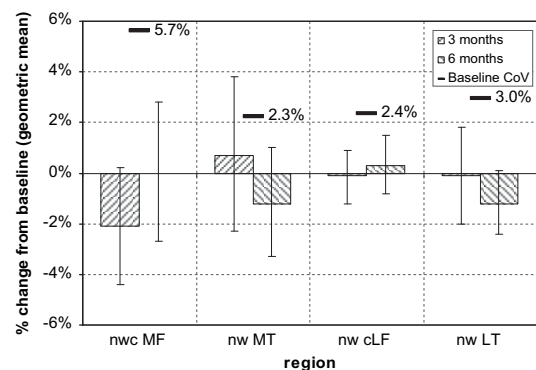


Fig. 3. Changes at 3 and 6 months in MRI cartilage thickness (ThCtAB) in nuclear (trimmed) central meniscal windowed regions on the medial and LT and femur. The mean and 95% CI estimates are from fitting ANOVA models to the log transformed data. The ANOVA models include the baseline measure as a covariate. The baseline is the geometric mean of the value at Visits 2a and 2b. CI s are not corrected for multiple comparisons. The baseline CoV overlaid is the within-subject.

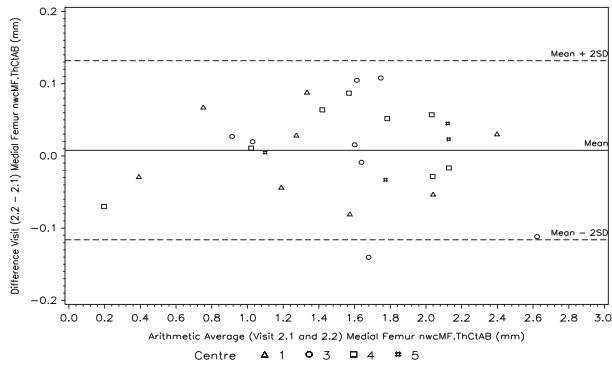


Fig. 4. Bland–Altman plot showing the agreement between the Visit 2a and Visit 2b values of the primary variable $nwcMF \cdot ThCtAB$.

MRI cartilage morphology variables were tested at 3 and 6 months, without consideration of the multiplicity of statistical comparisons, so this single significant (uncorrected) change should be interpreted with caution.

Baseline test–retest variabilities were low, and are comparable with other published MRI cartilage morphology reproducibilities^{5,14,29–32}, but were often higher than the average change seen over 3 and 6 months. The changes seen in this study are consistent with those reported over 12 months in similar cohorts from the OAI incidence group^{5,6}.

Prior studies using a sub-regional approach have found the greatest changes are in the central weight bearing portions of the medial TF joint—hence the reason we focused our attention here. The study by Pelletier *et al.* found the greatest change in the central medial tibia followed by central medial femur². Similarly Wirth *et al.* found that rate of cartilage loss was greater in central sub-regions than in entire FT cartilage plates.

Power calculations were performed to estimate the overall sample size required for a parallel group design to have 80% power to detect a complete halting of cartilage thinning (0% change) in a treatment group in comparison with a placebo group exhibiting the thinning shown in this study, where the change in both groups has the variability observed in this study, for a 10% significance level analyzed using a *t* test to compare the changes between groups, and allowing for a 10% drop-out. For the primary variable ($nwcMF \cdot ThCtAB$) over 3 months using a repeat mean baseline (where a change of -2.0% from mean baseline was observed with a variability of 7.14% in this study) an overall sample size of 353

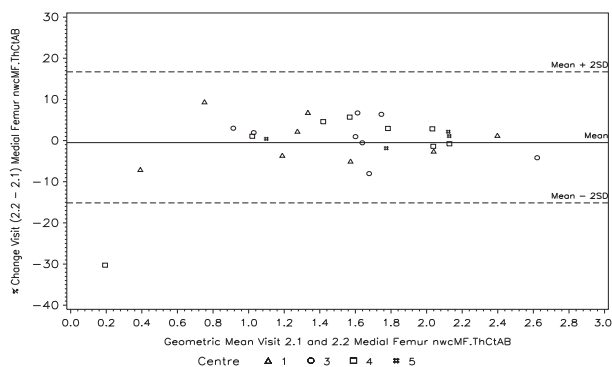


Fig. 5. Bland–Altman plot (calculated on the log-scale) showing agreement between the Visit 2a and Visit 2b values of the primary variable $nwcMF \cdot ThCtAB$ in terms of percentage change.

Table IV

Measurement repeatability for thickness and bone area: repeatability is shown for the difference in measurements between the Visit 2a and 2b images

ACRAC region	ThCtAB	tAB
	CoV (%)	CoV (%)
MF	2.2	1.3
nwcMF	5.7	1.2
MT	2.5	0.8
nwMT	2.3	1.5
LF	2.8	1.9
LT	3	1.7

would be required. For the primary variable ($nwcMF \cdot ThCtAB$) over 3 months using a single-baseline (where a change of -2.5% from first baseline was observed with a variability of 6.46% in this study) an overall sample size of 187 would be required. If the treatment group had less effect than the complete halting proposed here the sample sizes would be even larger.

However, given the lack of confidence in the mean changes observed due to the relatively high variabilities, in particular the fact that the mean changes in cartilage thickness are not significantly different from zero, power calculations based on the point estimates of the changes and variabilities observed in this study should be treated with extreme caution. The apparent improved power with single-baseline as opposed to double baseline may well be spurious. The actual sample size required for a 3-month study may be much greater. Indeed, from these data the possibility that there is no overall change in cartilage thickness during the first 3 months of such a study cannot be ruled out.

There are a number of limitations of this study. The size of the study sample investigated was small and could have contributed to the lack of meaningful and consistent change at both 3 and 6 months. There is the possibility that methods other than those based on the analysis of anatomical sub-regions may have provided better discrimination of subjects.

Taken alone, this study gives no confidence that measures of cartilage thickness obtained using the MR image acquisition and analysis deployed in this study provide a viable biomarker endpoint for future 3- or 6-month disease modification PoC studies in a highly stratified sample. Therefore cartilage thickness measured using MRI should not be considered as a biomarker for 3-month follow-up studies of OA without compelling additional data. Six-month follow-up studies may be possible, but given this study and currently available data this strategy should be considered high risk. Finally, this is an active area of research, and future developments in image acquisition, analysis and patient selection may yet yield viable imaging biomarkers for short, small PoC studies.

Conflict of interest

The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Acknowledgements

We would like to thank the participants and staff of the MeMO Osteoarthritis Study. We are grateful to the dedicated group of study coordinators whose skills were essential in assuring the successful conduct of this study. Central image coordination was done by VirtualScopics and the reading of plain radiographs was performed by Saara Totterman. Segmentations of the cartilage were performed by Bledwyn Woodward, with supervision by a musculoskeletal radiologist, Dr Charles Hutchinson of the University of Manchester. Dr Hunter is funded by an ARC Future Fellowship.

References

- Eckstein F, Burstein D, Link TM. Quantitative MRI of cartilage and bone: degenerative changes in osteoarthritis (Review). *NMR Biomed* 2006;19:822–54 (238 refs).
- Pelletier JP, Raynaud JP, Berthiaume MJ, Abram F, Choquette D, Haraoui B, et al. Risk factors associated with the loss of cartilage volume on weight-bearing areas in knee osteoarthritis patients assessed by quantitative magnetic resonance imaging: a longitudinal study. *Arthritis Res Ther* 2007;9:R74.
- Cicuttini FM, Wluka AE, Wang Y, Stuckey SL, Cicuttini FM, Wluka AE, et al. Longitudinal study of changes in tibial and femoral cartilage in knee osteoarthritis. *Arthritis Rheum* 2004;50:94–7 (See comment).
- Raynaud JP, Martel-Pelletier J, Berthiaume MJ, Labonte F, Beaudoin G, de Guise JA, et al. Quantitative magnetic resonance imaging evaluation of knee osteoarthritis progression over two years and correlation with clinical symptoms and radiologic changes. *Arthritis Rheum* 2004;50:476–87.
- Hunter DJ, Niu J, Zhang Y, Totterman S, Tamez J, Dabrowski C, et al. Change in cartilage morphometry: a sample of the progression cohort of the Osteoarthritis Initiative. *Ann Rheum Dis* 2009;68:349–56.
- Eckstein F, Maschek S, Wirth W, Wyman B, Hudelmaier M, Nevitt M, et al. One year change of knee cartilage morphology in the first release of participants from the Osteoarthritis Initiative Progression Subcohort – association with sex, body mass index, symptoms, and radiographic OA status. *Ann Rheum Dis* 2009;68:674–9.
- Williams T, Holmes AMR, Waterton J, Taylor C, Creamer P, Nash A. Cartilage loss in osteoarthritis detected by statistical shape analysis of magnetic resonance images (Abstract). *Osteoarthritis Cartilage* 2005;13(Suppl A):228.
- Felson DT. Obesity and osteoarthritis of the knee (Review). *Bull Rheum Dis* 1992;41:6–7 (10 refs).
- Garnero P, Ayrat X, Rousseau JC, Christgau S, Sandell LJ, Dougados M, et al. Uncoupling of type II collagen synthesis and degradation predicts progression of joint damage in patients with knee osteoarthritis. *Arthritis Rheum* 2002;46:2613–24 (Comment).
- Sharma L, Song J, Felson DT, Cahue S, Shamiyeh E, Dunlop DD. The role of knee alignment in disease progression and functional decline in knee osteoarthritis. *JAMA* 2001;286:188–95 (Erratum appears in *JAMA* 2001 Aug 15; 286 (7): 792.).
- Cicuttini F, Wluka A, Hankin J, Wang Y, Cicuttini F, Wluka A, et al. Longitudinal study of the relationship between knee angle and tibiofemoral cartilage volume in subjects with knee osteoarthritis. *Rheumatology* 2004;43:321–4.
- Teichtahl AJ, Davies-Tuck ML, Wluka AE, Jones G, Cicuttini FM. Change in knee angle influences the rate of medial tibial cartilage volume loss in knee osteoarthritis. *Osteoarthritis Cartilage* 2009;17:8–11.
- Eckstein F, Wirth W, Hudelmaier M, Stein V, Lengfelder V, Cahue S, et al. Patterns of femorotibial cartilage loss in knees with neutral, varus, and valgus alignment. *Arthritis Rheum* 2008;59:1563–70.
- Sharma L, Eckstein F, Song J, Guermazi A, Prasad P, Kapoor D, et al. Relationship of meniscal damage, meniscal extrusion, malalignment, and joint laxity to subsequent cartilage loss in osteoarthritic knees. *Arthritis Rheum* 2008;58:1716–26.
- Felson DT, McLaughlin S, Goggins J, LaValley MP, Gale ME, Totterman S, et al. Bone marrow edema and its relation to progression of knee osteoarthritis. *Ann Intern Med* 2003;139:330–6.
- Hunter DJ, Zhang YQ, Niu JB, Tu X, Amin S, Clancy M, et al. The association of meniscal pathologic changes with cartilage loss in symptomatic knee osteoarthritis. *Arthritis Rheum* 2006;54:795–801.
- Altman RD, Hochberg M, Murphy WAJ, Wolfe F, Lequesne M. Atlas of individual radiographic features in osteoarthritis. *Osteoarthritis Cartilage* 1995;3(Suppl A):3–70.
- Kraus VB, Vail TP, Worrell T, McDaniel G, Kraus VB, Vail TP, et al. A comparative assessment of alignment angle of the knee by radiographic and physical examination methods. *Arthritis Rheum* 2005;52:1730–5.
- Cooke TD, Sled EA, Scudamore RA. Frontal plane knee alignment: a call for standardized measurement. *J Rheumatol* 2007;34:1796–801.
- Peterfy C, Li J, Zaim S, Duryea J, Lynch J, 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. *Skeletal Radiol* 2003;32:128–32.
- Kellgren JH, Lawrence JS. Atlas of Standard Radiographs. Oxford: Blackwell Scientific; 1963.
- Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee (Review). *Osteoarthritis Cartilage* 2008;16:1433–41 (40 refs).
- Williams T, Taylor C, Gao Z, Waterton J. Corresponding articular cartilage thickness measurements in the knee joint by modelling the underlying bone. In: Ellis R, Peter T, Eds. Proceedings of the Seventh International Conference on Medical Image Computing & Computer Assisted Intervention. MICCAI 2003 2003. Heidelberg.
- Cootes T, Twining C, Petrovic V, Schestowitz R, Taylor C. Groupwise construction of appearance models using piecewise affine deformations. *Proc Br Mach Vis Conf* 2005;2:888.
- Williams T, Taylor C, Waterton J, Holmes A. Population analysis of knee cartilage thickness maps using model based correspondences. *Proc IEEE Int Symp Biomed Imaging* 2004;1:193–6.
- Eckstein F, Ateshian G, Burgkart R, Burstein D, Cicuttini F, Dardzinski B, et al. Proposal for a nomenclature for magnetic resonance imaging based measures of articular cartilage in osteoarthritis. *Osteoarthritis Cartilage* 2006;14:974–83.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
- 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.
- Hellio Le Graverand M, Buck R, Wyman B, Vignon E, Mazzuca S, Brandt K, et al. Change in regional cartilage morphology and joint space width in osteoarthritis participants versus healthy controls – a multicenter study using 3.0 Tesla MRI and Lyon Schuss radiography. *Ann Rheum Dis* 2008.
- Eckstein F, Buck RJ, Burstein D, Charles HC, Crim J, Hudelmaier M, et al. Precision of 3.0 Tesla quantitative magnetic resonance imaging of cartilage morphology in a multicentre clinical trial. *Ann Rheum Dis* 2008;67:1683–8.
- Kornaat PR, Reeder SB, Koo S, Brittain JH, Yu H, Andriacchi TP, et al. MR imaging of articular cartilage at 1.5 T and 3.0 T: comparison of SPGR and SSFP sequences. *Osteoarthritis Cartilage* 2005;13:338–44.
- Kornaat PR, Koo S, Andriacchi TP, Bloem JL, Gold GE. Comparison of quantitative cartilage measurements acquired on two 3.0 T MRI systems from different manufacturers. *J Magn Reson Imaging* 2006;23:770–3.