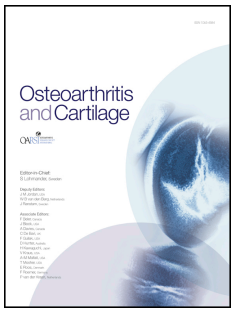
Accepted Manuscript



Automated Segmentation and Analysis of Normal and Osteoarthritic Knee Menisci from Magnetic Resonance Images - Data from the Osteoarthritis Initiative

Anthony Paproki, Craig Engstrom, Shekhar S. Chandra, Ales Neubert, Jurgen Fripp, Stuart Crozier

PII: S1063-4584(14)01152-2

DOI: 10.1016/j.joca.2014.06.029

Reference: YJOCA 3199

To appear in: *Osteoarthritis and Cartilage*

Received Date: 7 March 2014

Revised Date: 9 June 2014

Accepted Date: 28 June 2014

Please cite this article as: Paproki A, Engstrom C, Chandra SS, Neubert A, Fripp J, Crozier S, Automated Segmentation and Analysis of Normal and Osteoarthritic Knee Menisci from Magnetic Resonance Images - Data from the Osteoarthritis Initiative, *Osteoarthritis and Cartilage* (2014), doi: 10.1016/j.joca.2014.06.029.

Automated Segmentation and Analysis of Normal and Osteoarthritic Knee Menisci from Magnetic Resonance Images - Data from the Osteoarthritis Initiative

Anthony Paproki[1,2,*],
A.P, anthony.paproki@uqconnect.edu.au
School of Information Technology and Electrical Engineering,
The University of Queensland,
St Lucia, QLD 4027, Australia

Craig Engstrom[3],
C.E, c.engstrom@uq.edu.au

Shekhar S. Chandra[1],
S.S.C, shekhar.chandra@uq.edu.au

Ales Neubert[1,2],
A.N, ales.neubert@uqconnect.edu.au

Jurgen Fripp[1],
J.F, jurgen.fripp@csiro.au

Stuart Crozier[2],
S.C, stuart@itee.uq.edu.au

1. The Australian e-Health Research Centre, CSIRO Computational Informatics, Royal Brisbane and Women's Hospital, Herston QLD 4029, Australia

2. School of Information Technology and Electrical Engineering, The University of Queensland, St Lucia, QLD 4027, Australia

3. School of Human Movement Studies, The University of Queensland, St Lucia, QLD 4072, Australia

Abstract

Objective: To validate an automatic scheme for the segmentation and quantitative analysis of the medial (MM) and lateral meniscus (LM) in magnetic resonance (MR) images of the knee joint.

Method: We analysed sagittal water-excited dual-echo steady-state MR images of the knee joint from a subset of the Osteoarthritis Initiative cohort. The MM and LM were automatically segmented in the MR images based on a 3D deformable model approach. Quantitative parameters including volume, subluxation and tibial-coverage were automatically calculated from the segmentations for comparison (Wilcoxon tests) between knees with variable radiographic osteoarthritis (rOA), medial and lateral joint space narrowing (mJSN, lJSN) and pain characteristics. Automatic segmentations and estimated parameters were evaluated for accuracy using manual delineations of the menisci in 88 pathological knee MR examinations at baseline and 12 months time-points.

Results: The median (95% confidence-interval) Dice similarity index ($2 * |Auto \cap Manual|/(|Auto| + |Manual|) * 100$) between the manual and automated segmentations for the MM and LM were 78.3%(75.0—78.7), 83.9%(82.1—83.9) at baseline and 75.3%(72.8—76.9), 83.0%(81.6—83.5) at 12 months. Pearson coefficients between automatic and manual segmentation parameters ranged from *r*=0.70 to *r*=0.92. MM in rOA and mJSN knees had significantly greater subluxation and smaller tibial-coverage than no-rOA and no-mJSN knees. LM in rOA knees had significantly greater volumes and tibial-coverage than no-rOA knees.

Conclusion: Our automated method successfully segmented the menisci in normal and osteoarthritic knee MR images and detected meaningful morphological differences in the MM and LM with respect to rOA and JSN. Our approach will facilitate analyses of the menisci in prospective MR cohorts such as the OAI for investigations into pathophysiological changes occurring in early OA development.

Keywords: medial meniscus, lateral meniscus, automated segmentation, morphometric analysis, osteoarthritis, MRI

Automated Segmentation and Analysis of Normal and Osteoarthritic Knee Menisci from Magnetic Resonance Images - Data from the Osteoarthritis Initiative

Anthony Paproki[1,2,*],
A.P, anthony.paproki@uqconnect.edu.au
School of Information Technology and Electrical Engineering,
The University of Queensland,
St Lucia, QLD 4027, Australia

Craig Engstrom[3],
C.E, c.engstrom@uq.edu.au

Shekhar S. Chandra[1],
S.S.C, shekhar.chandra@uq.edu.au

Ales Neubert[1,2],
A.N, ales.neubert@uqconnect.edu.au

Jurgen Fripp[1],
J.F, jurgen.fripp@csiro.au

Stuart Crozier[2],
S.C, stuart@itee.uq.edu.au

1. The Australian e-Health Research Centre, CSIRO Computational Informatics, Royal Brisbane and Women's Hospital, Herston QLD 4029, Australia

2. School of Information Technology and Electrical Engineering, The University of Queensland, St Lucia, QLD 4027, Australia

3. School of Human Movement Studies, The University of Queensland, St Lucia, QLD 4072, Australia

Abstract

Objective: To validate an automatic scheme for the segmentation and quantitative analysis of the medial (MM) and lateral meniscus (LM) in magnetic resonance (MR) images of the knee.

Method: We analysed sagittal water-excited dual-echo steady-state MR images of the knee from a subset of the Osteoarthritis Initiative cohort. The MM and LM were automatically segmented in the MR images based on a deformable model approach. Quantitative parameters including volume, subluxation and tibial-coverage were automatically calculated for comparison (Wilcoxon tests) between knees with variable radiographic osteoarthritis (rOA), medial and lateral joint space narrowing (mJSN, lJSN) and pain. Automatic segmentations and estimated parameters were evaluated for accuracy using manual delineations of the menisci in 88 pathological knee MR examinations at baseline and 12 months time-points.

Results: The median (95% confidence-interval) Dice similarity index ($2 * |Auto \cap Manual|/(|Auto| + |Manual|) * 100$) between manual and automated segmentations for the MM and LM were 78.3%(75.0—78.7), 83.9%(82.1—83.9) at baseline and 75.3%(72.8—76.9), 83.0%(81.6—83.5) at 12 months. Pearson coefficients between automatic and manual segmentation parameters ranged from *r*=0.70 to *r*=0.92. MM in rOA/mJSN knees had significantly greater subluxation and smaller tibial-coverage than no-rOA/no-mJSN knees. LM in rOA knees had significantly greater volumes and tibial-coverage than no-rOA knees.

Conclusion: Our automated method successfully segmented the menisci in normal and osteoarthritic knee MR images and detected meaningful morphological differences with respect to rOA and JSN. Our approach will facilitate analyses of the menisci in prospective MR cohorts such as the OAI for investigations into pathophysiological changes occurring in early OA development.

Introduction

Quantitative analyses of the medial meniscus (MM) and lateral meniscus (LM) from three-dimensional (3D) magnetic resonance (MR) imaging offer opportunities to better understand the pathophysiological processes involved in the structural and functional degeneration of the menisci associated with osteoarthritis (OA)[1-3]. Recent semi[4-11] and fully-quantitative[8-15] MR studies have reported significant differences in the volume, tibial-coverage and subluxation of the menisci between knees with distinctive radiographic OA (rOA), medial and lateral joint space narrowing (mJSN, lJSN) or pain scores. While MR scoring methods provide good reproducibility and reliability for clinical evaluation of the menisci[4-6], acquisition of detailed quantitative data on these structures through MR segmentation offers increased measurement precision for investigating the *in-vivo* 3D morphological and biochemical characteristics of these fibro-cartilaginous discs (e.g. T2, T1ρ imaging[16-19], analysis of volume changes with OA or post surgery[20-21]).

Manual segmentation of the menisci from 3D MR images is a time- and expertise intensive process (35 minutes reported for segmentation of a single coronal water-excited double-echo steady-state (weDESS) MR[9]). Specifically, it requires numerous subjective interpretations for separating adjacent structures with comparable signal contrasts which predispose to low intra-rater reproducibility and high inter-rater variability[12]. A desirable direction is the automation the MR segmentation and analysis.

Several semi-automatic methods for the 3D segmentation of the menisci have been developed to reduce both analysis time and rater biases[19,20]. However these still require expert training and varying levels of manual intervention. In terms of fully automated segmentation approaches[22-25], good accuracy, as measured with the Dice similarity index (DSI)[26], has been achieved for the MM (75±10%) and LM volumes (77±10%)[24] and a total meniscal volume (81±3%)[25] although these methods were only validated on healthy menisci.

To the best of our knowledge, results and validation of fully automatic segmentations of the menisci from MR images of individuals with knee rOA have not been published. There are substantial technical challenges for automated segmentation of the menisci with pathological damage or degeneration which give rise to a spectrum of structural and biochemical tissue changes which, as illustrated in Fig. 1, are variably associated with increases in signal heterogeneity and shape variability[1,14-18]. Consequently, segmentation approaches that assume homogeneous signal intensity in the menisci are not well suited for morphometric analyses of the menisci in knees with rOA[19,20], and although methods that combined shape- and image-priors provided

promising leads[24,25], only preliminary results on the automatic segmentation of healthy meniscus have been reported in relatively small populations (N<14).

The objectives of this study were to 1) develop a fully-automatic method for the segmentation and quantitative analysis of the individual MM and LM from MR images of the knee, 2) quantitatively evaluate the accuracy of the automatic segmentation and estimations of derived parameters such as volume, subluxation and tibial coverage and 3) to explore the sensitivity of the method in the detection of meaningful changes in meniscus parameters across individuals with various rOA grades, mJSN, lJSN and pain.

[Figure 1] **/** [Table 1]

Material and Method

Patient and MR Image Datasets

The MR images used in this study were obtained from the Osteoarthritis Initiative (OAI) database, which is available for public access at http://www.oai.ucsf.edu/. Three Datasets (A), (B) and (C) of sagittal 3D weDESS MR images of the knee featuring a high spatial-resolution (0.37x0.37mm matrix, 0.7mm slice thickness) and signal-to-noise ratio well-suited for accurate morphological analyses of the meniscus, were selected from the OAI image release 0.E.1, 1.E.1, 3.E.1 and 5.E.1. Imaging protocol and knee positioning were standardized across all subjects[27]. Dataset (A) consisted of MR examinations of knee pathology from 88 patients selected from the OAI baseline and 12-month image releases. The MM and LM were manually segmented in all the MR examinations of Dataset (A) and kindly provided by Imorphics (Manchester, UK). The manual segmentations were performed by a single operator trained by a musculoskeletal radiologist (Charles Hutchinson) and an expert segmenter (Mike Bowes) and had passed the Imorphics cartilage segmentation training protocol, requiring an intra-observer coefficient of variation lower than 3% on paired test images. The segmentations were reviewed by the expert segmenter. These manual segmentations, which were performed blind to the present study, were used to train and validate of our automated segmentation algorithm. Datasets (B) and (C) consisted of 22 and 129 subjects (left and right knees) selected from the OAI Progression (definite rOA) and Incidence (asymptomatic with increased risks of developing OA) cohorts at baseline, 12, 24 and 36 months. Automated segmentations of the baseline MR images from Datasets (B) and (C) were undertaken for visual assessments of the performance of the segmentation method (results provided) and exploratory data analyses on meniscal volume, subluxation and tibial coverage in a larger cohort with a wider spectrum of healthy and

pathological meniscal morphologies. Additional results of automated analyses of the menisci in the longitudinal Datasets (B) and (C) are reported as supplementary material. Relevant demographics and clinical data are provided in Table 1.

[Figure 2]

Automatic Menisci Segmentation

The proposed automatic MR image segmentation method is based on a 3D active shape model (ASM) scheme[29] which involves deforming statistical shape models (SSMs) of the MM and LM in the MR image based on a template matching procedure[30,31].

A 3D SSM mathematically describes the direction and the magnitude of shape variability of a training-set of triangulated surfaces[29]. These models are characterised by a mean-shape which changes in a plausible manner (anatomically credible) based on a set of shape-parameters (illustrated in Supplementary Fig. S1.a). In ASM-fitting schemes, SSMs are frequently used to restrain the deformation from converging towards unlikely shapes deviating excessively from typical shapes of the training-set. In this work, three separate SSMs (1. combined menisci, 2. individual MM and 3. individual LM) were trained based on the manual segmentations of Dataset (A) at baseline. These SSMs were deformed in the MR images using image-feature models[30,31], which comprised 1D template intensity profiles typically surrounding the menisci in the training-set (illustrated in Supplementary Fig. S1.b). The ASM utilised these image-feature models in order to find the intensity profiles most similar to that of the template profiles in the new image to segment. Technical background regarding the generation of the models is provided in 'Supplementary Data A'.

The segmentation pipeline, detailed bellow and in Fig. 2.a, involved four steps: (1) image preprocessing, (2) ASM initialisation, (3) ASM-fitting and (4) post-processing. The method was implemented in C++ based on the Insight[32] and Visualisation Toolkits[33] (implementation details are provided in 'Supplementary Table S1').

I. MR Image Preprocessing

In this first stage, the MR image to segment – denoted I – was normalised to a fixed intensity range (0±200) using linear rescaling and preprocessed using a median smoothing algorithm (radius 1x1x1) in order to reduce the image noise and increase signal homogeneity within structures.

II. Affine initialisation

In the initialisation stage, an average menisci surface – denoted S_I – was aligned to a likely meniscus region of I based on the registration of an average knee image to I. Underlying methods utilised to generate the average knee image and menisci surface is described in 'Supplementary Data A'. The average knee image was first registered to I using an affine registration algorithm[34], and the obtained transformation was propagated to the average surface, resulting in a surface S_I approximately aligned with the meniscus region in I. To refine the initialisation, the meniscus region was extracted from both the average knee image and I (2mm around S_I), and the registration process was repeated with the cropped images.

For an individual with multiple time-point scans, the MR images were first co-registered and averaged into a subject-specific mean image using groupwise registration[35], and the mean-image obtained was used for the initialisation of all the time-points.

The initial pose and shape parameters of the ASM were then estimated from this obtained surface. An example of an initial segmentation obtained after this stage (by voxelising S_I into segmentation masks) is provided in Fig. 2.a(II).

III. Active Shape Model Fitting.

The SSMs were then deformed towards the most likely shape and position in I based on the template profile matching process illustrated in Fig. 2.a (shaded area) and described in detail elsewhere[30,31]. Summarising, for a given point k of S_I, a grey level profile $P_{I,k}$ longer than that of the image-feature model is extracted along the surface normal (positive and negative direction) and compared to the template profiles $P_{i,k}$ of the image feature model. As shown in Fig. 2.b, the template profiles $P_{i,k}$ are translated along the case profile $P_{I,k}$ and the normalised-cross-correlation $\gamma \in [0,1]$[30,31] is computed at each position. The translation offset of the profile $P_{i,k}$ which maximises γ is then used to translate the point k of S_I along its normal, thus deforming the surface. Once all the points of S_I have been translated, the deformed surface is restrained to a bounded space representative of typical menisci shapes by either constraining the shape-parameters of the SSM or smoothing the surface. The process is then iterated until the maximum number of iteration is reached

(supplementary Table S1).

Optimizing the ASM-fitting process for the segmentation of the menisci involved three parts. A combined SSM encoding the pose variability was deformed in I to refine the initial pose of the menisci. This step was performed using a 2 level Gaussian image-pyramid scheme to avoid converging towards local minima. In a second pass, individual SSMs of the MM and LM describing the local shape variability were separately deformed in I to obtain likely morphologies. SSMs were used to constrain the deformation during the 2 first stages of the fitting process[29].

To account for the shape variability not described by the SSMs and allow the ASMs to deform towards shapes slightly different than that of the training-set, a third pass deformed separate MM and LM ASMs in I without SSM constraints. Finally, smoothing was applied to remove noise from the deformed surface and the surface was voxelised to create the initial segmentation masks.

IV. Segmentation Post-Processing

To correct any small over-segmentation, a post processing classification method was applied to the menisci masks. As shown in Fig. 2.c, the tissue intensity properties of the MM and LM were estimated by a Gaussian distribution (mean μ, variance σ^2), and each voxel was assigned a probability of being meniscal tissue based upon its distance to μ. Since the intensity of meniscal tissues is expected to be lower than that of the surrounding articular cartilages in the weDESS MR images, voxels featuring intensities lower than $\mu+\sigma$ were classified as meniscal tissue and other voxels were discarded. To account for inherent signal intensity heterogeneity and tears within the menisci, potentially excluded from the meniscal tissue due to high signal intensity, “defects” in the internal portions of the image mask were marked as unclassified outliers and treated as meniscal tissue for quantitative analyses.

[Figure 3]

Quantitative Analysis

Based upon the segmentations and 3D reconstructions, the menisci were automatically analysed for volume, tibial coverage and subluxation parameters, which are often altered in individuals with knee rOA[7-11,13-15]. The volumes were computed by numerical integration of image-voxels belonging to the segmented menisci. The menisci subluxation and tibial-coverage parameters, illustrated in Fig. 3.a,b, required the identification of the

tibial bone and plateau (bone-cartilage interface), which were automatically obtained in each MR image following the method described by *Fripp et al.*[36]. The MM and LM coverage areas were calculated as the percentage of the medial and lateral tibial plateau surfaces (Fig. 3, yellow) covered by the individual menisci (Fig. 3, orange)[8,14]. The subluxation parameter was computed as the maximum distance between the external margin of the meniscus and that of the tibial plateau (Fig. 3.b green and red curves) when the meniscus position was 'external' relative to the tibial plateau, otherwise the minimum distance (signed negatively) was used[8,14].

Validation Strategy

The automated segmentation algorithm was applied to all 88 MR images of Dataset (A) with manual segmentations at baseline (V00) and 12 month (V01) time-points and quantitatively validated using a leave-one-out strategy (each case currently segmented was omitted from the training stage). The automatic and manual menisci segmentations were compared using the sensitivity, specificity, DSI[26] and mean absolute surface distance[37] (MASD) values as per Eq. 1:

$$\begin{aligned} Sensitivity &= TP/(TP+FN)*100 \\ Specificity &= TN/(TN+FP)*100 \\ DSI &= 2*|A\cap M|/(|A|+|M|)*100 \\ MASD &= (D(A,M)+D(M,A))/2) \end{aligned} \quad (1)$$

in which TP, TN, FP, FN are the number of true positives, true negatives, false positives and false negatives, and A and M are the automatic and manual segmentation masks respectively. The sensitivity, specificity, DSI and MASD quantified the percentage of true positives, true negatives, the spatial overlap and the average forward and backward Euclidean distances (D(x,y)) between automatic and manual segmentations[37]. For both the MM and LM, differences in DSI values were examined using Wilcoxon rank-sum tests across rOA grades and Wilcoxon signed-rank tests between time-points (significance-level 0.05). Non-parametric tests were used due to a negative skew in the DSI distributions.

Associations between meniscal parameters estimated from the automatic and manual segmentation data were investigated using the Pearson product-moment correlation coefficient[38], the intraclass correlation

coefficient (ICC - two-way random single measure)[39] and Bland-Altman analyses[40]. Coefficients above 0.75 were interpreted as good, while coefficients between 0.5 and <0.75 were interpreted as moderate. To account for outliers and the negative skew of the DSI distributions, the correlation analyses were performed on Dataset (A) trimmed by 5% of the DSI extrema.

Using the baseline imaging data pooled over all datasets, meniscal volume, subluxation and tibial coverage were compared for differences 1) between rOA groups (such that no(confirmed)-rOA = grade 0 or I, mild-rOA=grade II and advanced-rOA=grade III-IV), 2) between medial and lateral JSN groups (grades 0, I and II) and 3) between pain-score groups (WOMAC=0, 0<WOMAC<=10 and 10<WOMAC<=20) using Wilcoxon rank-sum tests adjusted for false discovery rate[41] (significance-level: 0.05).

All statistical analyses were performed using 'R 3.0'.

[Table 2]

Results

Segmentation Validation

There was good spatial overlap between the manual and automated segmentations of the MM (median(95%CI) DSI_{V00}=78.3%(75.0—78.7), DSI_{V01}=75.3%(72.8—76.9)) and LM volumes (DSI_{V00}=83.9%(82.1—83.9), DSI_{V01}=83.0%(81.6—83.5)) at both time points (Table 2). For each meniscus, there were no significant differences in DSI values across the rOA grades ($p>0.05$ for all Wilcoxon rank-sum tests) and between V00 and V01 (Wilcoxon signed-rank tests $p>0.05$). Segmentations for the MM and LM in representative cases corresponding to the interquartile mean, maximum and minimum DSI are provided in Fig. 4.a,b to visualise the typically good spatial overlap between the automatic and manual approaches. Severe damage to either or both of the menisci, as shown for the MM in Fig. 4.c, resulted in segmentation difficulties and low DSI values (<=60%) in a small number of cases (15/176≈8.5% for MM and 3/176≈1.7% for LM).

There were strong or moderate correlations (Fig. 5) between the manual and automated meniscal parameters at both V00 and V01 for the MM (r_{V00}=0.80, ICC_{V00}=0.80; r_{V01}=0.78, ICC_{V01}=0.78) and LM (r_{V00}=0.91, ICC_{V00}=0.90; r_{V01}=0.89, ICC_{V01}=0.88) volume, the MM (r_{V00}=0.83, ICC_{V00}=0.83; r_{V01}=0.70, ICC_{V01}=0.69) and LM (r_{V00}=0.92, ICC_{V00}=0.91; r_{V01}=0.89, ICC_{V01}=0.89) subluxation and the MM (r_{V00}=0.82, ICC_{V00}=0.81; r_{V01}=0.81, ICC_{V01}=0.79) and LM (r_{V00}=0.83, ICC_{V00}=0.82; r_{V01}=0.71, ICC_{V01}=0.70) tibial coverage.

Comparisons between the manual and automated volume data using Bland-Altman plots showed for both the MM and LM an even distribution of the differences between methods across the range of meniscal measures (no apparent funnelling effects) with a bias of (-4.45%, 6.46%), (-0.525mm, -0.266mm) and (-1.98%, -1.64%) for the (MM, LM) volume, subluxation and tibial coverage, respectively. Automatic segmentation and quantitative analysis results obtained for each patient of Datasets (A), (B) and (C) can be publicly accessed online at http://milxview.csiro.au/msk_meniscus/xplorer_studies/Public[42]. Observations across automated segmentations of both the MM and LM of Datasets (B) and (C) were visually comparable to those obtained and evaluated for (A), indicating an overall robustness of the method.

[Figure 4/5]

Quantitative Analysis

As reported in Table 3, in rOA and mJSN knees, the MM had significantly more subluxation and less tibial coverage than no-rOA/no-mJSN knees. So did the MM of advanced-rOA knees compared to mild-rOA knees. mJSN and advanced-rOA knees also had significantly greater MM volume than no-mJSN and no-rOA knees respectively. The subluxation of the MM was significantly greater in knees with advanced-rOA compared to mild-rOA knees.

For the LM, knees with rOA had significantly greater meniscal volume and tibial-coverage than no-rOA/no-lJSN knees. The volume of the LM was also greater in knees with lJSN.

No significant differences were noted between different groups of pain score in any of the meniscal parameters.

Automated segmentations of the MM and LM were also performed to obtain volume, subluxation and tibial coverage measures at baseline, 12, 24 and 36 months follow-up for the OAI Progression and Incidence datasets. At this stage descriptive data, as reported in Supplementary Table S2, have been generated with additional data input such as rOA grade and compartmental JSN progression required for downstream analyses and validation.

[Table 3]

Computational Time

All the experiments were performed on a dual 6-core Intel Xeon Westmere X5670 (2.93GHz) workstation. Using our fully-automatic method, the mean±SD CPU time required to segment the MM and LM from an

individual MR examination was 27.2±1.8 minutes (min=24.3, max=32.3), and the time required to perform the quantitative analysis was 2.4±0.2 minutes (min=2.1, max=3.4) minutes.

[Figure 6]

Discussion

This study is the first to successfully provide automatic segmentation and quantitative analysis of both the MM and LM from MR images of individuals with knee rOA. Leave-one-out experiments showed good spatial agreement between the manual and automated segmentations of the individual MM and LM, with overall median DSI values of 77.1% and 83.5% in analyses of knee MR examinations for individuals presenting a large spectrum of JSN, sclerotic bone and osteophytes (rOA grades II-IV). For both the MM and LM in individuals with knee rOA, our scheme may provide a good alternative to the semi-automatic method of *Swanson et al.*[20] which performed direct segmentations of T2-Maps (more challenging to segment) and achieved a DSI of 69%. In terms of DSIs, our approach compared favourably to previous automated segmentation approaches of these structures in healthy states[24] and although the mean DSI (81.9%) obtained by *Zhang et al.*[25] slightly outperforms our current results, a direct comparison is difficult since the segmentation of individual menisci was not reported.

The automatic estimation of quantitative parameters was sufficiently accurate ($0.70<r<0.92$) to discern meaningful cross-sectional differences in the volume, tibial-coverage and subluxation of the meniscus between groups with variable rOA characteristics. In particular, the MM showed overall greater subluxation and smaller tibial-coverage area in individuals with rOA and mJSN, concurrent with recent findings[8,13-15]. The LM showed a greater median volume in knees with rOA and mJSN. The tibial coverage of the LM was found significantly greater in knees with rOA (4.4% relative difference, which corroborate results from a recent study although this difference was not reported significant[15]), but was not significantly different between knees with and without lJSN.

The good DSI values and successful identification of significant differences between groups, particularly in relation with the meniscal subluxation and tibial-coverage (eg. 0.43mm and 2.0% absolute difference detected in subluxation and tibial-coverage of the MM between rOA and no-rOA knees) which have been associated with cartilage loss[3,8], suggest that the present method would be suitable to efficiently analyse and monitor the evolution meniscus morphological characteristics with OA development in large populations. The automatic method also provides opportunities for investigations into the biochemical changes of the meniscus with OA,

requiring accurate co-registration schemes (eg. *Xue et al.*[43]) to align the high-resolution MR image with the biochemical MR sequence (eg. T1ρ, T2, dGEMERIC MR). However, as with all automated methods, quality control procedures are required to detect the small number of segmentation failures. Our initial experience found web applications[42] (e.g. http://milxview.csiro.au/msk_meniscus/xplorer_studies/Public) to be efficient to perform this task, although further investigations are required to ensure their effectiveness.

The primary advantages of our method are: (1) it does not require any manual MR image processing, (2) it provides good segmentation of the MM and LM as separate labels, (3) it performs well on knees with rOA and visual inspections showed equivalent performance in healthy knees, (4) it readily segments torn menisci and finally (5) it does not require prior identification of the bones or articular cartilages within the knee.

There are some limitations with the present research. First, the method was only evaluated on weDESS MR images acquired as part of the OAI. Further validation is required for to assess the applicability of the method on clinically focused sequences such as intermediately-weighted 2D fast-spin-echo (FSE) and 3D-FSE.

Regarding the performance of the method on the OAI weDESS MR images, another possible limitation of the method was the decrease DSI values with rOA severity (for MM) and between time-points (Table 2). The primary reason for these differences relates to the increase in meniscus shape complexity and MR signal heterogeneity associated with disease progression, which blurred the boundaries with articular-cartilages and weakened the features driving the ASM. These differences were not significant (Table 2: $p>0.05$ for all comparisons), suggesting that the method maintained reliable segmentations in the majority of the cases. Training the models of the segmentation algorithm using V01 yielded equivalent results, with a non-significant decrease in DSI values between time-points (Wilcoxon signed-rank test MM: $p=0.11$, $r=0.12$; LM: 0.53, $r=0.05$), which reduced the likelihood that this difference was induced by a strong time-point or training bias. In comparison with the segmentation of the LM, analyses of the anatomically more mobile MM presented greater challenges and a lower median DSI value was obtained. The primary cause of this was the greater variability of shapes and MR tissue-contrasts encountered in the MR images for this structure as a result of a more substantial expression of structural and biochemical alterations with OA. A median DSI of 77% (overall) still compares favourably with existing analyses of this structure in healthy states[24], highlighting the potential of the current segmentation approach for automated analyses of pathological menisci. The tip of the horns and the peripheral margins mid-way along the MM and LM were the areas that segmented least accurately (Fig. 6). These results stem from the unclear demarcations between the meniscal horns and ligaments and between the peripheral edges of the meniscus and fat (Fig. 1). From the high specificity and comparatively low

sensitivity reported in Table 2, we concluded that under-segmentation was the most common segmentation error obtained.

Several cases such as the MM shown in Fig. 4.c exhibited severe tissue destruction and our automated method failed in this specific instance of very advanced tissue loss. Our experience showed that these failed segmentations could be easily detected from the web applications previously mentioned, and with a failure rate (DSI<=60%) of 8.5% for MM and 1.7% for LM, we consider the method suitable for analyses of the menisci in a framework of early OA assessment.

In conclusion, our automated scheme is well suited to efficiently process and analyse large prospective MR cohorts, thereby presenting opportunities to facilitate epidemiological and interventional studies into morphological changes of the meniscus. The proposed method provides good accuracy for segmentation of the MM and LM meniscus from weDESS MR images from individuals with variably severe knee rOA (overall median DSI of 77.1% for MM and 83.5% for LM). Subsequent quantitative analyses obtained Pearson correlations ranging from 0.70 to 0.92 between manual and automatic volume, subluxation and tibial-coverage of the meniscus. Cross-sectional comparisons of the MM and LM parameters from various rOA and compartmental JSN groups provided results that corroborated previous manual findings.

Author's Contribution

All listed authors provided substantial contributions in (1) the conception and design of the study, or acquisition of data, or analysis and interpretation of data. All authors (2) participated in the redaction and revision of the paper. Finally, all authors (3) approved the final version of the article submitted.

In particular, AP, JF, and SC participated in the conception and design of the study. AP, SSC, AN, and JF implemented substantial parts of the method presented. JF, CE, and SC obtained the funding resources for the project. And finally, AP, CE, JF, and SC provided helpful background, statistical expertise and clinical insights crucial for the research study.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors would like to thank Mike Bowes (Imorphics) and Dr Charles Hutchinson (University of Manchester) for kindly providing the manual segmentations for public access.

Role of the funding sources

This research was supported under Australian Research Council's Linkage Projects funding scheme LP100200422.

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 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.

References

1. Englund M, Guermazi A, Lohmander LS. The Meniscus in Knee Osteoarthritis. Rheumatic Disease Clinics of North America 2009;35(3):579–90.
2. Chan W, Lang P, Stevens M, Sack K, Majumdar S, Stoller D, et al. Osteoarthritis of the knee: comparison of radiography, CT, and MR imaging to assess extent and severity. American Journal of Roentgenology 1991;157(4):799–806.
3. 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 and rheumatism 2006;54(3):795-801.

4. Peterfy C.G, Guermazi A, Zaim S, Tirman P.F., Miaux Y, White D et al., Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis, Osteoarthritis and Cartilage 2004;12(3):177–90.
5. 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(2):206–11.
6. Hunter D.J, Guermazi A, Lo G.H, Grainger A.J, Conaghan P.G, Boudreau R.M, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score), Osteoarthritis and Cartilage 2011;19(8):990-1002.
7. Berthiaume MJ, Raynauld JP, Martel-Pelletier J, Labonté F, Beaudoin G, Bloch Da, et al. Meniscal tear and extrusion are strongly associated with progression of symptomatic knee osteoarthritis as assessed by quantitative magnetic resonance imaging. Annals of the Rheumatic Diseases 2005;64(4):556–63.
8. Bloecker K, Guermazi A, Wirth W, Benichou O, Kwoh CK, Hunter DJ, et al. Tibial coverage, meniscus position, size and damage in knees discordant for joint space narrowing - data from the Osteoarthritis Initiative. Osteoarthritis and Cartilage 2013;21(3):419–27.
9. Bloecker K, Guermazi A, Wirth W, Kwoh CK, Resch H, Hunter D.J, et al. Correlation of semiquantitative vs quantitative MRI meniscus measures in osteoarthritic knees: results from the Osteoarthritis Initiative. Skeletal radiology 2014;43(2):227–32.
10. Am Jung K, Lee SC, Hwang SH, Yang KH, Kim DH, Sohn JH, et al. High frequency of meniscal hypertrophy in persons with advanced varus knee osteoarthritis. Rheumatology international 2010;30(10):1325–33.
11. Hwang S, Jung K, Lee W, Yang K, Lee D, Carter A, et al. Morphological changes of the lateral meniscus in end-stage lateral compartment osteoarthritis of the knee. Osteoarthritis and Cartilage 2012;20(2):110–6.
12. Siorpaes K, Wenger A, Bloecker K, Wirth W, Hudelmaier M, Eckstein F. Interobserver reproducibility of quantitative meniscus analysis using coronal multiplanar DESS and IWTSE MR imaging. Magnetic Resonance in Medicine 2012;67(5):1419–26.
13. Bloecker K, Wirth W, Hunter DJ, Duryea J, Guermazi A, Kwoh, CK, et al. Contribution of regional 3D meniscus and cartilage morphometry by MRI to joint space width in fixed flexion knee radiography--a between-knee comparison in subjects with unilateral joint space narrowing. European journal of radiology 2013;82(12):832–9.

14. Wirth W, Frobell RB, Souza RB, Li X, Wyman BT, Le Graverand MPH, et al. A three-dimensional quantitative method to measure meniscus shape, position, and signal intensity using MR images: A pilot study and preliminary results in knee osteoarthritis. Magnetic Resonance in Medicine 2010;63(5):1162–71.
15. Wenger A, Wirth W, Hudelmaier M, Noebauer-Huhmann I, Trattnig S, Bloecker K, et al. Meniscus Body Position, Size, and Shape in Persons With and Persons Without Radiographic Knee Osteoarthritis: Quantitative Analyses of Knee Magnetic Resonance Images From the Osteoarthritis Initiative. Arthritis & Rheumatism 2013;65(7):1804–11.
16. Li X, Pai A, Blumenkrantz G, Carballido-Gamio J, Link T, Ma B, et al. Spatial distribution and relationship of T1ρ and T2 relaxation times in knee cartilage with osteoarthritis. Magnetic Resonance in Medicine 2009;61(6):1310–8.
17. Mayerhoefer ME, Welsch GH, Riegler G, Mamisch TC, Materka A, Weber M, et al. Feasibility of texture analysis for the assessment of biochemical changes in meniscal tissue on T1 maps calculated from delayed gadolinium-enhanced magnetic resonance imaging of cartilage data: comparison with conventional relaxation time measurements. Investigative Radiology 2010;45(9):543–7.
18. Wang L, Chang G, Xu J, Vieira RLR, Krasnokutsky S, Abramson S, et al. T1ρ MRI of menisci and cartilage in patients with osteoarthritis at 3T. European Journal of Radiology 2012;81(9):2329–36.
19. Rauscher I, Stahl R, Cheng J, Li X, Huber MB, Luke A, et al. Meniscal Measurements of T1ρ and T2 at MR Imaging in Healthy Subjects and Patients with Osteoarthritis. Radiology 2008;249(2):591–600.18.
20. Swanson M, Prescott J, Best T, Powell K, Jackson R, Haq F, et al. Semi-automated segmentation to assess the lateral meniscus in normal and osteoarthritic knees. Osteoarthritis and Cartilage 2010;18(3):344–53.
21. Bowers ME, Tung GA, Fleming BC, Crisco JJ, Rey J. Quantification of meniscal volume by segmentation of 3 T magnetic resonance images. Journal of Biomechanics 2007;40(12):2811–5.
22. Sasaki T, Hata Y, Ando Y, Ishikawa M, Ishikawa H. Fuzzy rule-based approach to segment the menisci regions from MR images. In: Medical Imaging'99. International Society for Optics and Photonics 1999; 258–65.
23. Tamez-Pena JG, Totterman S, Parker KJ. Unsupervised statistical segmentation of multispectral volumetric MRI images. In: Medical Imaging'99. International Society for Optics and Photonics 1999; 300–11.

24. Fripp J, Bourgeat P, Engstrom C, Ourselin S, Crozier S, Salvado O. Automated segmentation of the menisci from MR images. In: IEEE International Symposium on Biomedical Imaging 2009; 510–3.
25. Zhang K, Lu W, Marziliano P. The unified extreme learning machines and discriminative random fields for automatic knee cartilage and meniscus segmentation from multi-contrast MR images. Machine Vision and Applications 2013;24(7):1459–72.
26. Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26(3):297–302.
27. Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. Osteoarthritis and Cartilage 2008;16(12):1433–41.
28. Kellgren J, Lawrence J. Radiological assessment of osteo-arthrosis. Annals of the Rheumatic Diseases 1957;16(4):494–502 .
29. Cootes TF, Taylor CJ, Cooper DH, Graham J. Active Shape Models-Their Training and Application. Computer Vision and Image Understanding 1995;61(1):38–59.
30. Chandra S, Dowling J, Shen K, Raniga P, Pluim JF, Greer PB, et al. Patient specific prostate segmentation in 3-D magnetic resonance images. IEEE Transactions on Medical Imaging 2012;31:1955–64.
31. Neubert A, Fripp J, Engstrom C, Schwarz R, Lauer L, Salvado O, et al. Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Physics in Medicine and Biology 2012;57(24):8357.
32. Ibanez L, Schroeder W, Ng L, Cates J. The ITK software guide: the insight segmentation and registration toolkit. Kitware Inc 2003;5.
33. Schroeder W, Martin K, Lorensen B. An Object-Oriented Approach To 3D Graphics. Prentice Hall 1997.
34. Rivest-Hénault D, Dowson N, Greer P, Dowling J. Symmetric affine registration of CT and MR pelvis scans for MR-based planning and adaptive prostate radiation therapy (Accepted). IOP Journal of Physics: Conference Series 2013.
35. Reuter M, Schmansky JN, Rosas HD, Fischl B. Within-subject template estimation for unbiased longitudinal image analysis. NeuroImage 2012;61(4):1402–18.
36. Fripp J, Crozier S, Warfield SK, Ourselin S. Automatic segmentation of the bone and extraction of the bone-cartilage interface from magnetic resonance images of the knee. Physics in Medicine and Biology 2007;52(6):1617–31.

37. Gerig G, Jomier M, Chakos M. Valmet: A new validation tool for assessing and improving 3D object segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2001. vol. 2208 of Lecture Notes in Computer Science. Springer. Springer Berlin Heidelberg 2001; 516–23.
38. Rodgers JL, Nicewander WA. Thirteen ways to look at the correlation coefficient. American Statistician 1988;4:59–66.
39. Koch GG. Intraclass correlation coefficient. Encyclopedia of Statistical Sciences 1983;4:212–7.
40. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet 1986;327(8476):307–10.
41. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 1995;57(1):289–300.
42. Bourgeat P, Dore V, Villemagne VL, Rowe CC, Salvado O, Fripp J. MilxXplore: a web-based system to explore large imaging datasets. Journal of the American Medical Informatics Association 2013;20(6):1046–52.
43. Xue N, Doellinger M, Fripp J, Ho CP, Surowiec RK, Schwarz R. Automatic model-based semantic registration of multimodal MRI knee data. Journal of Magnetic Resonance Imaging 2014.
44. Vapnik V. The nature of statistical learning theory. Springer 2000.
45. Breiman L. Random forests. Machine learning 2001;45(1):5–32.
46. Fripp J, Crozier S, Warfield SK, Ourselin S. Automatic Segmentation and Quantitative Analysis of the Articular Cartilages From Magnetic Resonance Images of the Knee. IEEE Transactions on Medical Imaging 2010;29(1):55–64.
47. Lorensen WE, Cline HE. Marching cubes: A high resolution 3D surface construction algorithm. In: Proceedings of the 14th annual conference on Computer graphics and interactive techniques. SIGGRAPH '87. New York, NY, USA 1987; 163–9.
48. Combes B, Prima S. An Efficient EM-ICP Algorithm for Symmetric Consistent Non-linear Registration of Point Sets. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010. vol. 6362 of Lecture Notes in Computer Science. Springer Berlin / Heidelberg 2010. 594–601.
49. Gower JC. Generalized procrustes analysis. Psychometrika 1975;40(1):33–51.
50. Jolliffe IT. Principal Component Analysis. 2nd ed. Springer-Verlag, New York 2002.

Figure legends

Figure 1: (a) Manual segmentation of the menisci in a 3D weDESS MR image acquired in the sagittal plane (patient 9056363, female, age 57, height 168.5cm, BMI 31.8kg/m^2, rOA grade III). *(left)* Coronal view, MM=medial meniscus, LM = lateral meniscus, FM=femur, T=tibia, C=cartilage, F=fat. *(Right)* Axial view, AH=anterior horn, PH = posterior horn. **(b)** A 3D sagittal weDESS MR image of healthy menisci demonstrating high tissue intensity homogeneity and clear demarcation between the surrounding cartilage and fat tissues in *(left)* coronal and *(right)* axial views (patient 9323403, male, age 51, height 161.8cm, BMI 27.4, rOA grade 0). **(c)** The menisci in a patient with moderate/severe rOA of the knee joint demonstrating "lesions" in the menisci in *(left)* coronal and *(right)* axial views (patient 9800677, male, age 65, 184.7cm, BMI 31.1kg/m^2, rOA grade III).

Figure 2: (a) Segmentation method flow diagram (axial view illustration; case 9056363, rOA grade III) demonstrating processing of the MM (orange) and LM (green) after affine initialisation, combined menisci ASM pose estimation, constrained MM and LM ASM fitting, MM and LM relaxation and tissue classification. Although shown in 2D, segmentation occurs in 3D. **(b)** Grey level profile matching. For a surface point *k* and associated profile $P_{I,k}$ of length 2*rL*+1 (with *r =1.5 a* padding ratio allowing the extraction of profiles larger than that of the image-feature-model), the template profiles $P_{i,k}$ of length 2*L*+1 are translated along $P_{I,k}$ and γ[30,31] is computed for each position. The profile and displacement maximising γ describe the displacement of the point *k* along its normal. The green line at the centre of each profile represents the menisci surface. **(c)** Post-processing stage for a MM with over-segmentation (green arrow) and a tear (blue arrow) visualised in the axial view. (1) Segmentation of the MM following ASM-fitting stage. (2) Tissue probability estimation within the MM (darker shades of blue denote a lower probability of meniscal tissue). (3) Tissue classification based on probability estimation. (4) Final segmentation of the MM following dilation and erosion, which allowed closure of the defect associated with the high signal intensity of the tear.

Figure 3: Schematic representation of the computation of the tibial-coverage and subluxation parameters. **(a)** A 3D rendering of the MM and LM (displayed as semi-transparent surfaces) tibial coverage areas (orange MM.Cov, LM.Cov) on the medial and lateral tibial plateau (yellow MM.TA and LM.TA). **(b)** Computation of the subluxation for the MM. The red (t_{ext}) and green (m_{ext}) points are the outermost points of the tibial plateau

(MM.TA) and the MM which, in this case, maximise the subluxation. The distance between these two points defines the subluxation parameter (blue arrow). A: anterior, P: posterior.

Figure 4: Qualitative assessment between manual (green overlay) and automatic (blue overlay) meniscal segmentations viewed as per right knee. *(left)* 3 axial slices focused on the MM, *(middle)* manual segmentation, *(right)* automatic segmentation. In **(a)**, from top to bottom, MM segmentation in cases situated at the interquartile mean (Case 9651690, DSI=77.6%, rOA grade III), interquartile minimum (Case 9602703, DSI=71.6%, rOA grade III), and interquartile maximum (Case 9954040, DSI=81.9%, rOA grade III). Similarly, in **(b)**, from top to bottom, LM segmentation in cases situated at the interquartile mean (Case 9382271, DSI=83.4%, rOA grade II), interquartile minimum (Case 9368622, DSI=80.6%, rOA grade IV), and interquartile maximum (Case 9698705, DSI=85.8%, rOA grade III). **(c)** is an illustration of segmentation failure caused by severe truncation of the MM (Case 9311328, DSI=37.0%, rOA grade III).

Figure 5: (a), **(b)** and **(c)** present the correlation and Bland-Atlman analyses performed for the MM *(left, green)* and LM *(right, blue)* volume, subluxation and tibial coverage parameters. The scatter-plots present the automatic segmentation parameters against the manual segmentation parameters, and the Bland-Altman analyses present the relative (for volume and tibial coverage) or absolute (for the subluxation) difference between automatic segmentation parameters and the manual segmentation parameters. The absolute error (expressed in mm) is used for the subluxation due to the presence of zero valued parameters.

Figure 6: Mapping of the median Hausdorff Distance (maximum forward/backward distance between automatic and manual surface)[37] between deformed ASM surfaces and the manual segmentations onto the menisci mean shape at V00 *(top)* and V01 *(bottom)* (all subjects from Dataset (A)). The blue and red areas characterised the smallest and largest distances to the manual segmentations. The tip of the horns and the external surface of the mid-compartment of both menisci were the areas most problematic to segment.

Supplementary Figure S1: (a) The major mode of variation (strongest λ_k of C) for the combined menisci SSM, MM SSM, and LM SSM, explained between -3 and +3 standard deviations from the mean-shape. **(b)** An example of N grey level profiles extracted along the surface normal vectors for a given surface. In this

example, the profiles extracted along the surface normals had a length L of 12. **(c)** The average knee image and surface used as atlas in the affine registration (initialisation).

Table 1: Demographic data of subjects analysed from the 3 datasets used in the present study. Readings for the rOA grades and JSN were based on site readings performed by a certified radiologist. rOA grades (near Kellgren/Lawrence grade[28]) at baseline are defined such that 0 = normal knee, 1 = not confirmed rOA, 2 = definite-mild rOA, 3 = moderate rOA, 4 = severe rOA. mJSN and lJSN are defined such that a grade of 0 = no-JSN (equivalent to Osteoarthritis Research Society International (OARSI) JSN grades 0), 1 = mild-JSN (equivalent to OARSI grades I-II) and 2 = severe-JSN (equivalent to OARSI grades III). The pain score is defined as the Western Ontario and McMaster Universities Arthritis Index (WOMAC).

	A		**B**		**C**	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
N	45	43	12	10	43	86
Age (yrs)	62.02±10.89	60.42±8.982	56.58±9.29	60.44±9.68	56.47±9.03	57.81±8.6
Height (cm)	176.7±6.39	163.1±5.80	178.3±5.68	164.6±7.56	176.9±5.85	162.2±6.46
Mass (kg)	96.31±14.76	83.67±14.87	94.54±18.04	78.14±16.36	86.24±14.27	75.24±14.46
BMI (kg/m^2)	30.51±3.87	31.65±5.26	29.64±4.97	28.68±4.98	27.55±4.22	28.59±5.27
Pain score ([0,20])	5.07±3.85	5.84±4.27	1.75±2.34	2.67±3.12	3.95±4.32	4.57±4.37
Time-points	2		4*		4*	
Left and Right	No		Yes		Yes**	
# Knees baseline	88		42		127	
# Total Knees	176		158		758	
rOA Grade (0;1;2;3;4)	(0, 0, 15, 56, 17)		(0, 9, 17, 14, 2)		(203, 33, 10, 7, 1)	
m.JSN score (0;1;2)	(16, 55, 17)		(30, 10, 2)		(240, 12, 2)	
l.JSN score (0;1;2)	(74, 14, 0)		(26, 16, 0)		(251, 3, 0)	

*30 and 104 patient knees were available at 4 Time-Points in (B) and (C), other patient were missing time-points
** Except for 3 cases

1

1

Table 2: Evaluation of the accuracy (median(MD), 95% confidence-interval(CI)) of the automated segmentation algorithm for the MM and LM volumes after the affine initialisation, ASM-fitting and classification (final) stages for the MM and LM volumes at V00 and V01. Final segmentation results are reported for the overall population and per OA grade.

V00	Sensitivity (%)		Specificity (%)		DSI (%)		MASD (MM)		
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI	(p, r)*
Medial Meniscus (V00)									
Affine	51.1	49.1-52.8	99.97	99.97-99.97	58.5	54.8-58.7	0.92	0.90-1.00	
ASM-fitting	77.8	75.5-79.3	99.97	99.97-99.98	77.6	74.0-77.9	0.51	0.49-0.61	
Final (overall)	**77.1**	**74.7-78.4**	**99.98**	**99.97-99.98**	**78.3**	**75.0-78.7**	**0.49**	**0.46-0.58**	
- OA Grade II	72.5	63.9-78.1	99.99	99.98-100.00	79.2	73.6-81.7	0.41	0.34-0.69	
- OA Grade III	78.0	75.8-79.5	99.98	99.97-99.98	78.7	74.4-79.2	0.5	0.46-0.61	*p=0.72, r=0.04*
- OA Grade IV	78.3	68.6-81.3	99.97	99.96-99.98	76.9	69.5-79.3	0.51	0.46-0.70	*p=0.25, r=0.21*
Lateral Meniscus (V00)									
Affine	51.0	50.4-55.4	99.97	99.97-99.97	58.9	56.2-60.7	0.86	0.82-0.94	
ASM-fitting	81.2	79.5-81.9	99.99	99.98-99.99	83.1	81.6-83.5	0.33	0.33-0.39	
Final (overall)	**79.0**	**77.7-80.1**	**99.99**	**99.99-99.99**	**83.9**	**82.1-83.9**	**0.33**	**0.32-0.38**	
- OA Grade II	75.5	71.9-78.9	99.99	99.99-99.99	82.2	79.5-84.7	0.32	0.29-0.41	
- OA Grade III	79.6	78.4-80.9	99.99	99.99-99.99	84.0	82.4-84.3	0.32	0.31-0.38	*p=0.49 ,r=0.08*
- OA Grade IV	81.1	73.1-83.1	99.99	99.98-99.99	84.1	78.1-85.2	0.35	0.29-0.73	*p=0.88, r=0.03*
V01	Sensitivity (%)		Specificity (%)		DSI (%)		MASD (MM)		
	MD	95% CI	MD	95% CI	MD	95% CI	MD	95% CI	(p, r)*
Medial Meniscus (V01)									
Affine	47.3	46.9-50.9	99.97	99.96-99.97	54.5	52.2-56.3	0.97	0.95-1.06	
ASM-fitting	78.0	76.0-79.2	99.97	99.96-99.97	74.2	71.6-75.8	0.58	0.55-0.67	
Final (overall)	**77.1**	**75.2-78.4**	**99.98**	**99.97-99.98**	**75.3**	**72.8-76.9**	**0.54**	**0.52-0.64**	p=0.07, r=0.14
- OA Grade II	75.7	68.8-78.6	99.99	99.97-99.99	81.5	71.3-83.3	0.38	0.33-0.71	
- OA Grade III	77.4	75.2-79.5	99.97	99.96-99.97	75.1	71.6-76.6	0.58	0.53-0.69	*p=0.07,r=0.21*
- OA Grade IV	78.5	71.9-79.8	99.97	99.96-99.98	74.8	68.9-78.6	0.54	0.47-0.72	*p=0.13,r=0.27*
Lateral Meniscus (V01)									
Affine	49.2	48.0-51.6	99.97	99.96-99.97	55.6	53.5-57.0	0.94	0.90-0.99	
ASM-fitting	81.3	79.1-81.6	99.99	99.98-99.99	82.8	80.9-83.0	0.34	0.35-0.41	
Final (overall)	**78.9**	**77.4-79.9**	**99.99**	**99.99-99.99**	**83.0**	**81.6-83.5**	**0.33**	**0.33-0.40**	p=0.18, r=0.10
- OA Grade II	75.3	70.3-79.3	99.99	99.99-99.99	82.4	79.1-84.8	0.33	0.30-0.47	
- OA Grade III	80.4	78.1-80.9	99.99	99.98-99.99	83.0	81.6-83.8	0.32	0.32-0.40	*p=0.79,r=0.02*
- OA Grade IV	78.8	72.3-81.3	99.99	99.98-99.99	83.2	74.6-84.7	0.36	0.31-0.72	*p=0.94,r=0.02*

*p-values and effect-size given in italics are the results of the Wilcoxon sum-rank tests between the DSI values of OA grade II verse the DSI values of OA grade III and IV. Underlined values given at time-point V01 correspond to the Wilcoxon signed-rank test results between the DSI values obtained at V00 and V01.

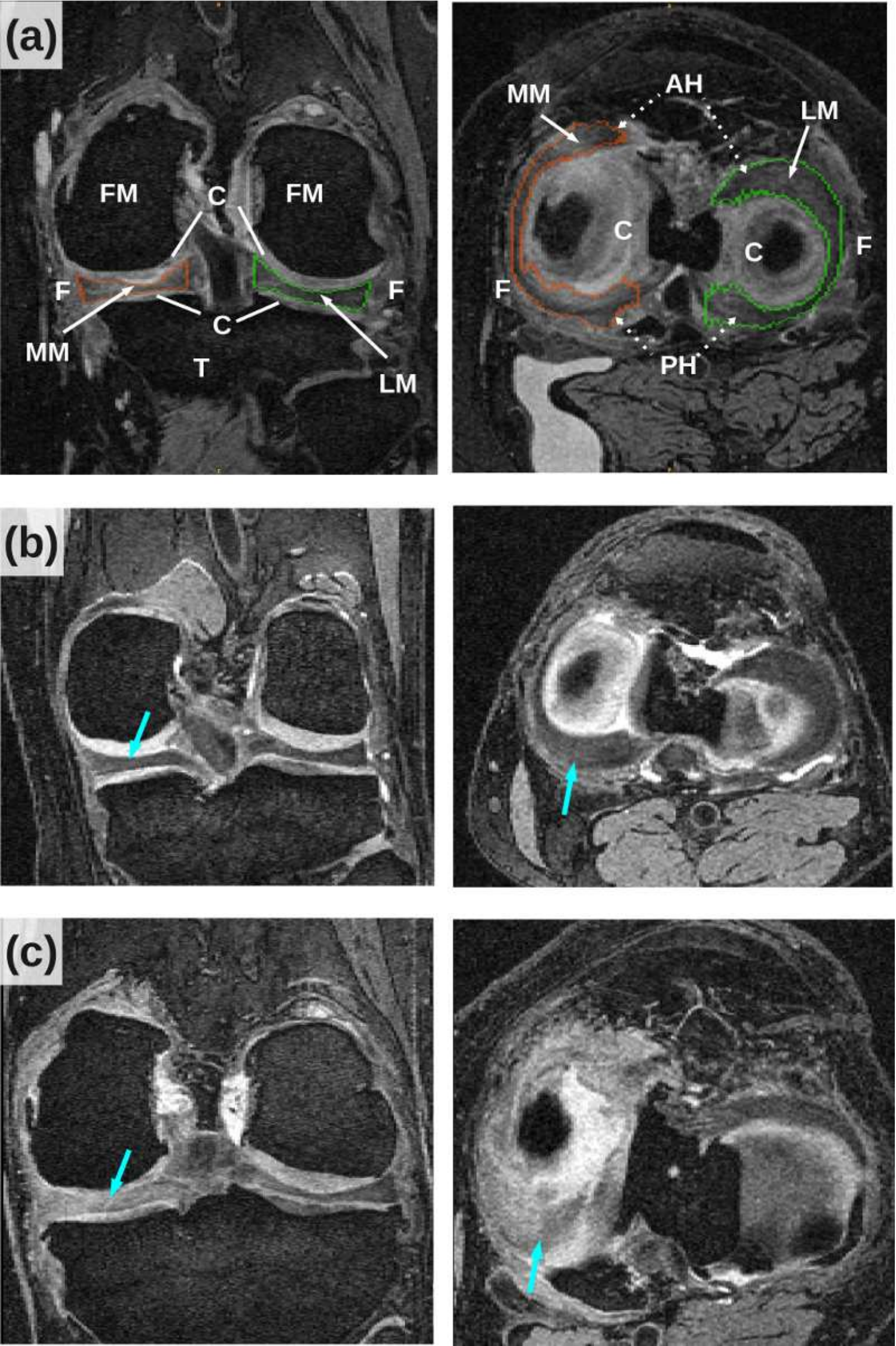
2
3

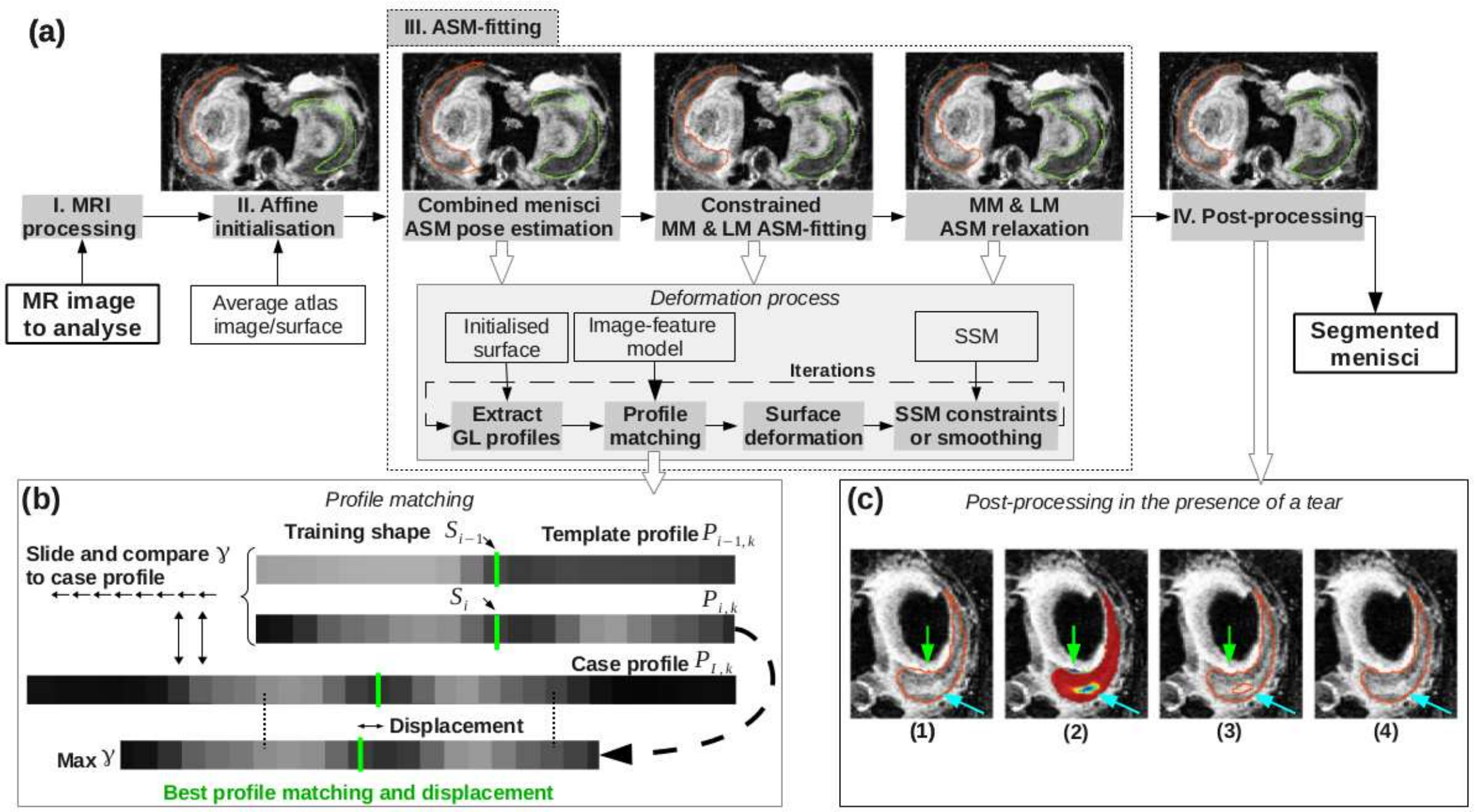
Table 3: Median values (MD), interquartile range (IQR), significance values and effect-sizes for MM and LM volume, subluxation, and tibial-coverage comparisons between knees with no-rOA, mild-rOA, and advanced-rOA, between knees with no-JSN, mild-JSN and severe-JSN and between 3 groups of patients with increasing WOMAC scores ([0],]0;10] and]10;20]). For the LM, only knees with no-JSN or mild-JSN were available. MM.Vol and LM.Vol are expressed in mm^3, MM.Sub and LM.Sub are expressed in mm and MM.Cov and LM.Cov are pressed in %.

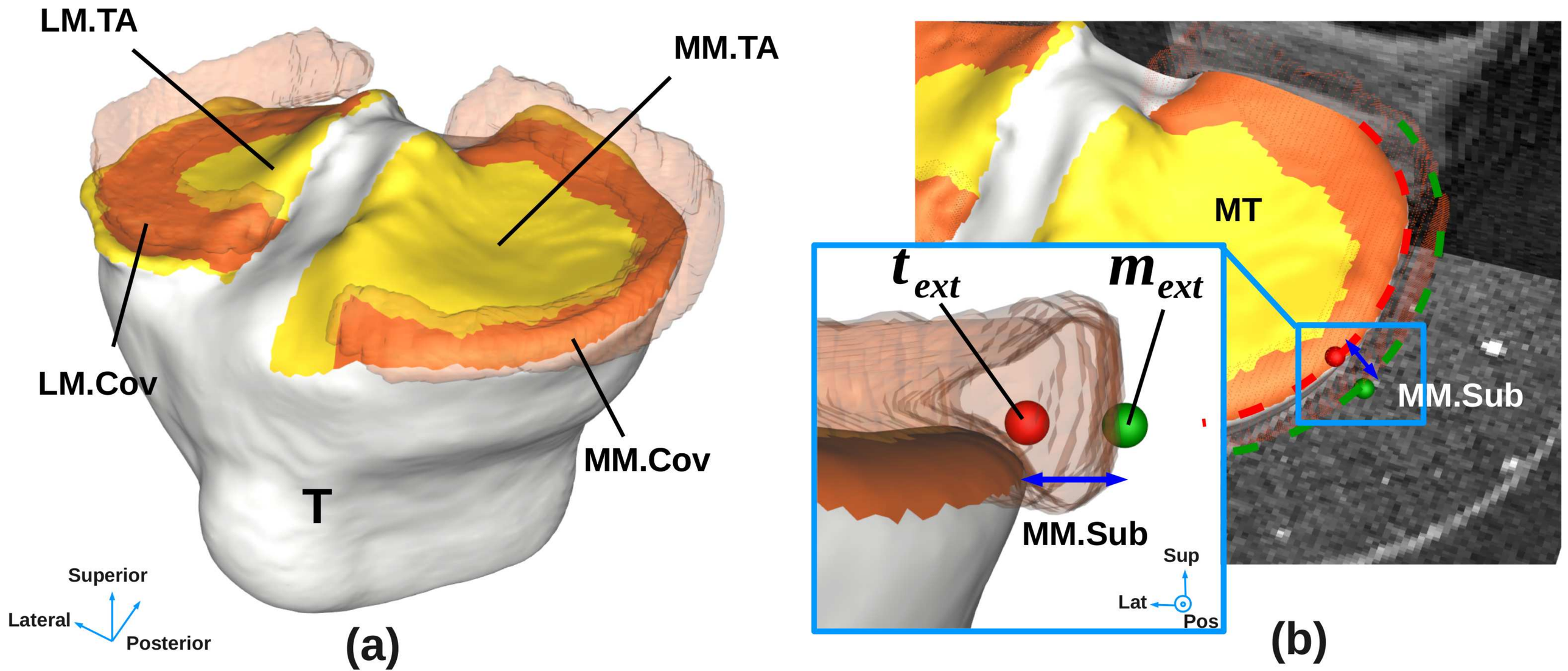
	Radiographic OA								
	no-rOA		**mild-rOA**		**advanced-rOA**		**p-value; effect-size**		
	MD	**IQR**	**MD**	**IQR**	**MD**	**IQR**	**no-rOA vs mild-rOA**	**no-rOA vs advanced-rOA**	**mild-rOA vs advanced-rOA**
MM.Vol	1949	1465-2406	2100	1484-2678	2350	1873-2857	0.126;0.09	**<0.001;0.27**	0.091,0.16
LM.Vol	1631	1386-2016	2331	1599-2854	2243	1730-2746	**<0.001;0.28**	**<0.001;0.40**	0.814,0.02
MM.Sub	2.31	1.31-3.38	2.74	2.10-3.94	4.59	3.56-5.50	**0.013;0.15**	**<0.001;0.54**	**<0.001;0.47**
LM.Sub	0.17	-0.14-0.94	0.57	0.00-1.58	0.60	0.00-1.31	0.097;0.11	0.078,0.12	0.902,0.01
MM.Cov	45.2	42.0-48.7	42.8	39.2-46.3	38.1	35.0-43.5	**0.016;0.14**	**<0.001;0.43**	**0.001,0.29**
LM.Cov	42.9	39.5-45.7	44.8	42.1-48.1	44.9	42.2-48.2	**0.014;0.15**	**<0.001;0.20**	0.865,0.015
	Medial and Lateral JSN								
	no-JSN		**mild-JSN**		**severe-JSN**		**p-value; r-value**		
	MD	**IQR**	**MD**	**IQR**	**MD**	**IQR**	**no-JSN vs mild-JSN**	**no-JSN vs severe-JSN**	**mild-JSN vs severe-JSN**
MM.Vol	1958	1472-2419	2295	1873-2825	2675	1934-3093	**<0.001;0.22**	**0.001,0.19**	0.290,0.11
LM.Vol	1751	1449-2243	2629	1723-2887	-	-	**<0.001;0.22**	-	-
MM.Sub	2.35	1.46-3.42	4.41	3.31-5.51	4.87	4.25-5.43	**<0.001;0.44**	**<0.001;0.37**	0.098,0.17
LM.Sub	0.25	-0.11-1.02	0.86	0.00-1.44	-	-	0.053;0.10	-	-
MM.Cov	44.5	41.3-48.3	39.3	36.5-44.8	35.3	29.8-41.1	**<0.001;0.31**	**<0.001;0.29**	**0.013,0.25**
LM.Cov	43.6	40.1-46.4	45.5	40.3-48.8	-	-	0.078;0.09	-	-
	WOMAC Score								
	0 (n=77)		**]0;10] (n=273)**		**]10;20] (n=34)**		**p-value; r-value**		
	MD	**IQR**	**MD**	**IQR**	**MD**	**IQR**	**0 vs]0;10]**	**0 vs]10;20]**	**]0;10] vs]10;20]**
MM.Vol	2017	1560-2439	2039	1502-2585	2120	1550-2543	0.822;0.03	0.822,0.05	0.822,0.01
LM.Vol	1810	1524-2260	1752	1447-2357	1975	1462-2273	0.887;0.03	0.887,0.01	0.887,0.03
MM.Sub	2.72	1.52-3.93	2.77	1.70-4.24	3.00	2.02-4.31	0.256;0.07	0.256,0.13	0.551,0.03
LM.Sub	0.42	0.00-1.06	0.19	-0.12-1.01	0.70	-0.082-1.84	0.173;0.09	0.476,0.07	0.173,0.09
MM.Cov	44.8	41.5-48.2	43.4	38.7-47.8	43.4	41.2-46.6	0.106;0.11	0.476,0.10	0.626,0.03
LM.Cov	43.9	41.3-46.6	43.5	39.9-46.4	44.0	39.5-46.8	0.501;0.07	0.875,0.02	0.761,0.04

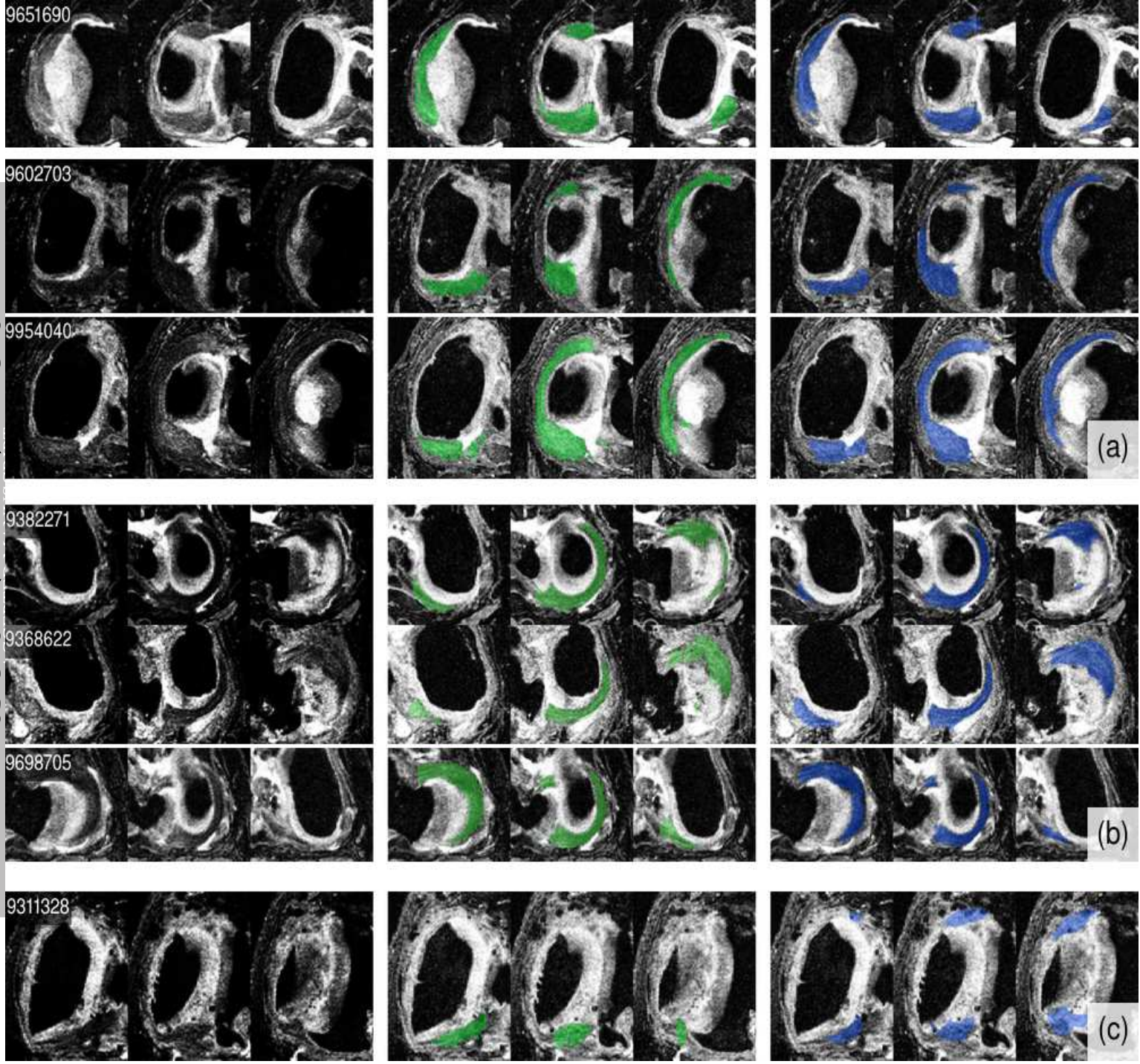
no-rOA = rOA grade 0 or 1, mild-rOA = grade II, advanced-rOA = grade III-IV; MM and LM parameters were tested against medial and lateral JSN respectively, with grade 0 = no-JSN, 1 = mild-JSN, 2 = severe-JSN. Differences between groups were tested using Wilcoxon rank-sum tests, with a significance level p=0.05. P-values were adjusted for multiple comparisons using false discovery rate[41].

1

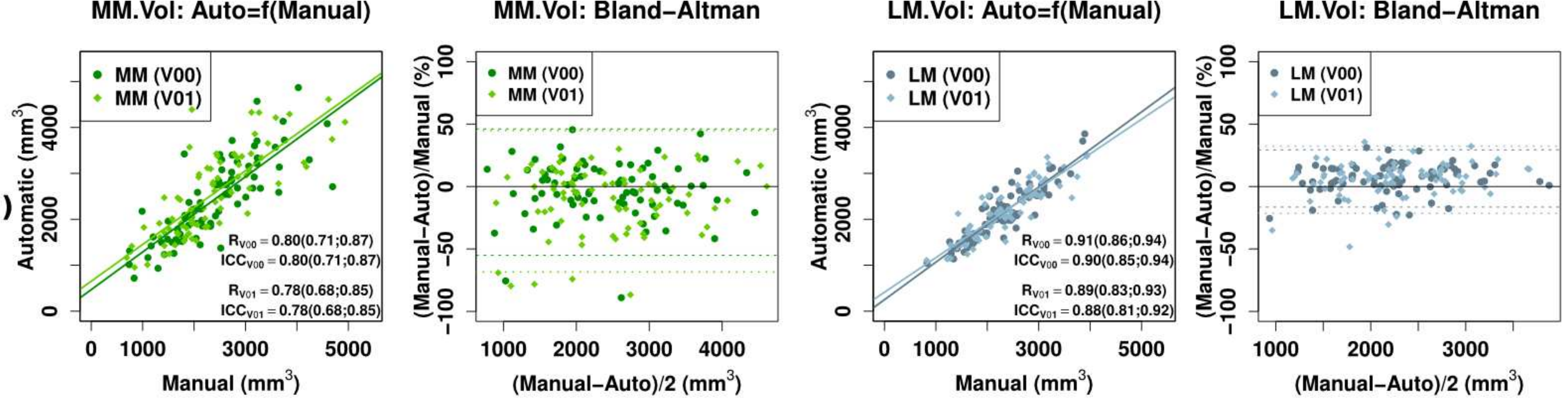

(a)
FM
C
FM
F
F
C
MM
T
LM
MM
AH
LM
C
C
F
F
PH
(b)
(c)


(a)
III. ASM-fitting
I. MRI processing
II. Affine initialisation
Combined menisci ASM pose estimation
Constrained MM & LM ASM-fitting
MM & LM ASM relaxation
IV. Post-processing
MR image to analyse
Average atlas image/surface
Segmented menisci
Deformation process
Initialised surface
Image-feature model
SSM
Iterations
Extract GL profiles
Profile matching
Surface deformation
SSM constraints or smoothing
(b)
Profile matching
Training shape S_{i-1}
Template profile $P_{i-1,k}$
Slide and compare γ to case profile
S_i
$P_{i,k}$
Case profile $P_{I,k}$
Displacement
Max γ
Best profile matching and displacement
(c)
Post-processing in the presence of a tear
(1)
(2)
(3)
(4)

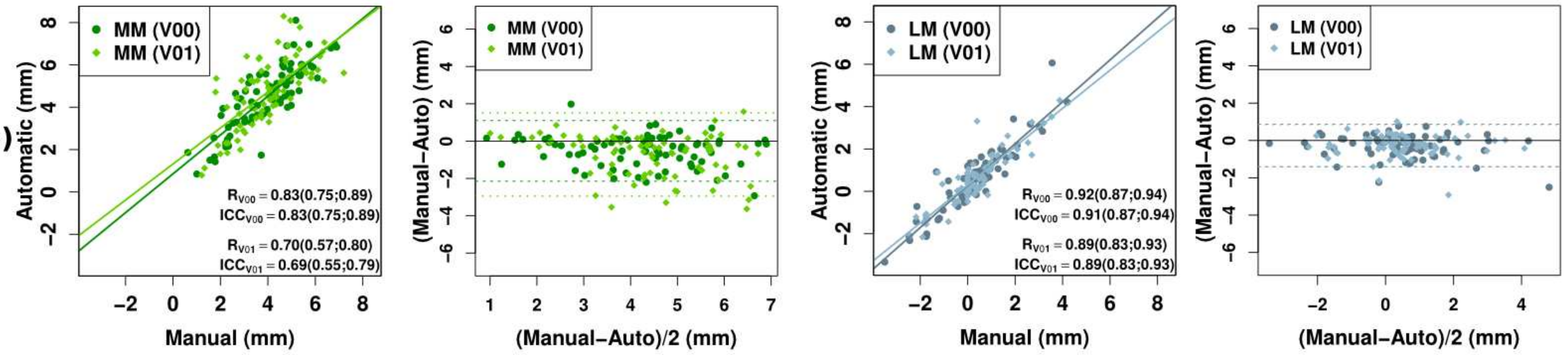

LM.TA
MM.TA
LM.Cov
MM.Cov
T
Superior
Lateral
Posterior
(a)
MT
t_{ext}
m_{ext}
MM.Sub
MM.Sub
Sup
Lat
Pos
(b)


9651690
9602703
9954040
9382271
9368622
9698705
9311328
(a)
(b)
(c)
Anterior
Posterior
Medial
Lateral

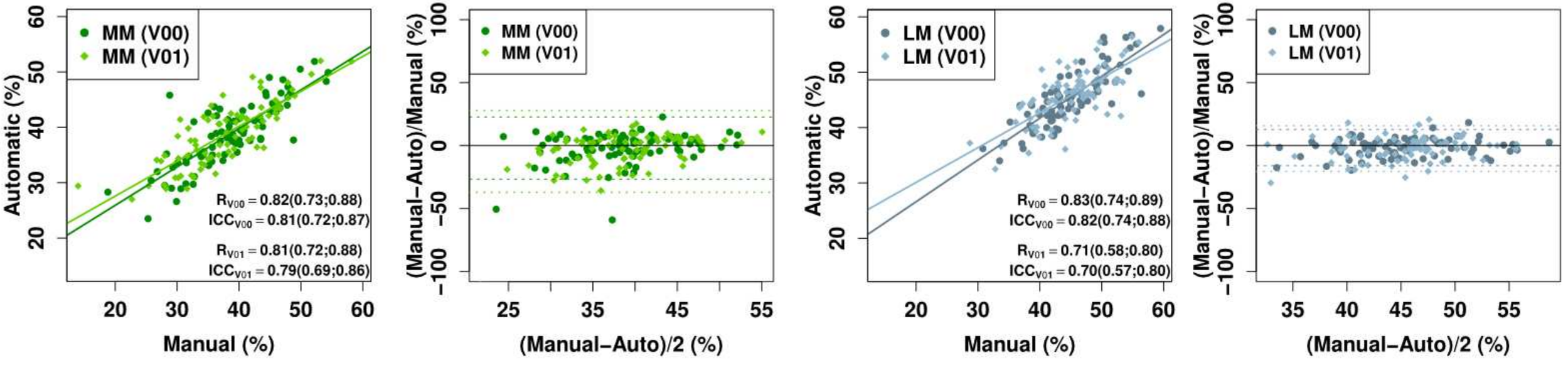
(a)


MM.Vol: Auto=f(Manual)
MM (V00)
MM (V01)
Automatic (mm3)
Manual (mm3)
RV00 = 0.80(0.71;0.87)
ICCV00 = 0.80(0.71;0.87)
RV01 = 0.78(0.68;0.85)
ICCV01 = 0.78(0.68;0.85)
MM.Vol: Bland–Altman
(Manual–Auto)/Manual (%)
(Manual–Auto)/2 (mm3)
LM.Vol: Auto=f(Manual)
LM (V00)
LM (V01)
RV00 = 0.91(0.86;0.94)
ICCV00 = 0.90(0.85;0.94)
RV01 = 0.89(0.83;0.93)
ICCV01 = 0.88(0.81;0.92)
LM.Vol: Bland–Altman

(b)


MM.Sub: Auto=f(Manual)
MM (V00)
MM (V01)
Automatic (mm)
Manual (mm)
RV00 = 0.83(0.75;0.89)
ICCV00 = 0.83(0.75;0.89)
RV01 = 0.70(0.57;0.80)
ICCV01 = 0.69(0.55;0.79)
MM.Sub: Bland–Altman
(Manual–Auto) (mm)
(Manual–Auto)/2 (mm)
LM.Sub: Auto=f(Manual)
LM (V00)
LM (V01)
RV00 = 0.92(0.87;0.94)
ICCV00 = 0.91(0.87;0.94)
RV01 = 0.89(0.83;0.93)
ICCV01 = 0.89(0.83;0.93)
LM.Sub: Bland–Altman

(c)


MM.Cov: Auto=f(Manual)
MM (V00)
MM (V01)
Automatic (%)
Manual (%)
RV00 = 0.82(0.73;0.88)
ICCV00 = 0.81(0.72;0.87)
RV01 = 0.81(0.72;0.88)
ICCV01 = 0.79(0.69;0.86)
MM.Cov: Bland–Altman
(Manual–Auto)/Manual (%)
(Manual–Auto)/2 (%)
LM.Cov: Auto=f(Manual)
LM (V00)
LM (V01)
RV00 = 0.83(0.74;0.89)
ICCV00 = 0.82(0.74;0.88)
RV01 = 0.71(0.58;0.80)
ICCV01 = 0.70(0.57;0.80)
LM.Cov: Bland–Altman

Superior
Lateral
Anterior

V00

Distance
1.914
1.435
0.957
0.478
0.000

Superior
Anterior
Lateral

V01

Superior
Lateral
Anterior

View: Lateral - Medial

View: Medial - Lateral

Supplementary Data C: Longitudinal results

Table S2. Median and interquartile range (IQR) of the automatic parameters computed for the MM and LM at baseline, 12, 24 and 36 months follow-up for the OAI Progression and Incidence datasets.

Medial Meniscus								
	Baseline		**12 Months**		**24 Months**		**36 Months**	
	Median	**IQR**	**Median**	**IQR**	**Median**	**IQR**	**Median**	**IQR**
				Volume (mm^3)				
OAI Progression (B)	2291	1870-2629	2551	2012-3244	2379	2053-2906	2405	2065-2899
OAI Incidence (C)	1884	1477-2346	1832	1395-2258	1804	1432-2279	1837	1496-2313
				Subluxation (mm)				
OAI Progression (B)	3.26	2.40-4.13	3.51	2.37-4.61	3.77	2.29-4.59	3.94	2.62-4.77
OAI Incidence (C)	1.96	1.17-3.10	2.04	1.24-2.88	2.19	1.22-3.07	2.12	1.15-2.98
				Tibial Coverage (%)				
OAI Progression (B)	39.6	37.1-46.0	43	37.9-46.1	42.8	35.0-46.5	38.2	36.9-45.2
OAI Incidence (C)	45.5	41.9-49.5	45.4	42.1-48.8	45.6	41.2-48.5	45.8	41.7-48.4

Lateral Meniscus								
	Baseline		**12 Months**		**24 Months**		**36 Months**	
	Median	**IQR**	**Median**	**IQR**	**Median**	**IQR**	**Median**	**IQR**
				Volume (mm^3)				
OAI Progression (B)	3054	2707-3542	3093	2865-3456	3139	2888-3536	3095	2846-3463
OAI Incidence (C)	1524	1347-1918	1534	1334-1974	1542	1319-1944	1609	1353-2002
				Subluxation (mm)				
OAI Progression (B)	0.82	0.094-1.48	0.86	0.29-1.73	0.91	0.22-1.37	0.76	0.00-1.67
OAI Incidence (C)	0.25	-0.097-0.92	0.19	-0.17-1.22	0.36	-0.029-1.15	0.2	-0.043-1.18
				Tibial Coverage (%)				
OAI Progression (B)	45.7	42.2-48.6	46.3	41.8-48.3	46	43.6-49.7	44.4	41.7-47.9
OAI Incidence (C)	43	40.0-45.5	43.2	40.2-46.5	42.9	40.2-46.8	43.7	39.9-46.2

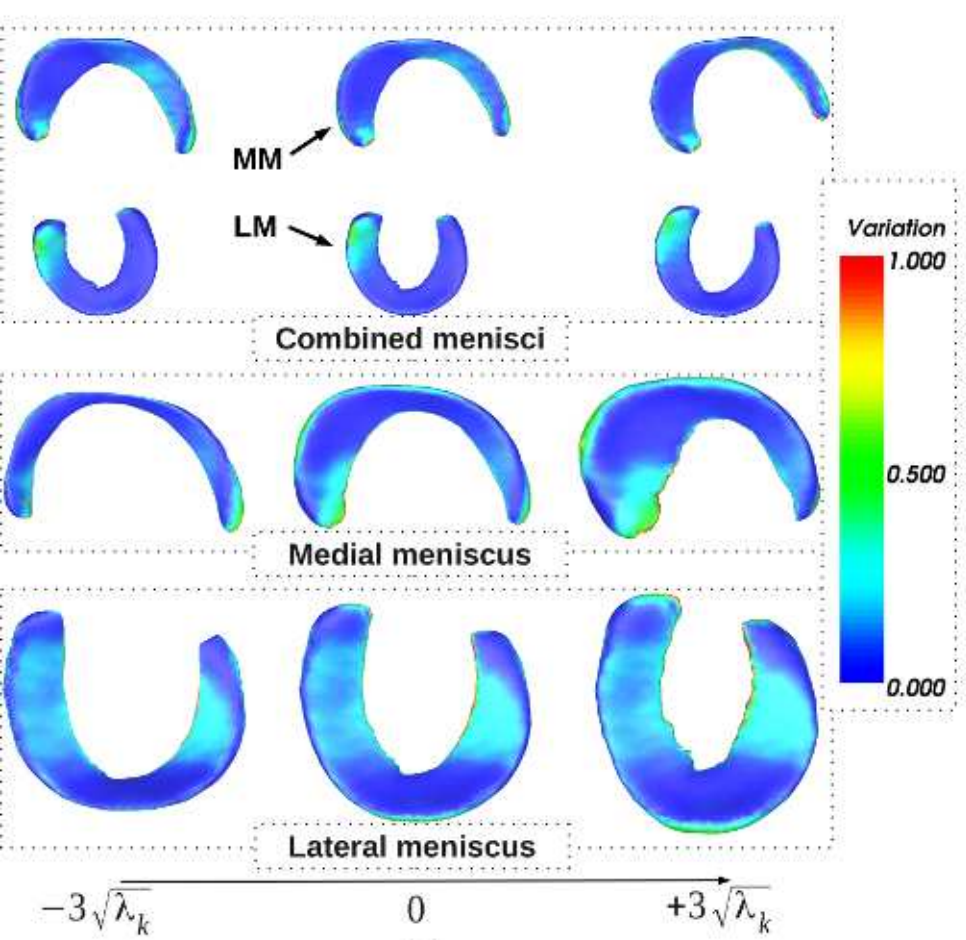
Supplementary Data B: Algorithm Parameters

Table S1. ASM-fitting implementation parameters. The segmentation of all subjects utilised the same parameters, which have been tuned based on training and observations.

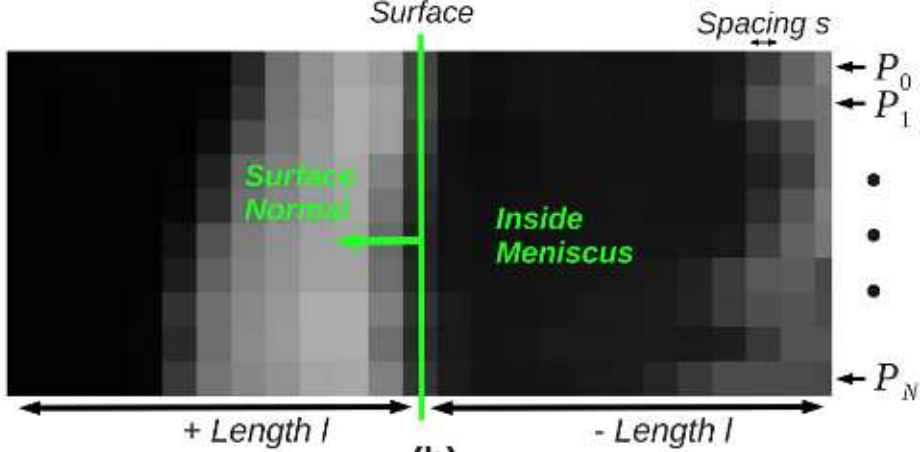
Fitting stage	Description	SSM # of modes	SSM deviation from $\bar{S}$	Profile length (2L+1)	Profile spacing	Fitting Iterations	Constraint type
(1) Combined menisci ASM-fitting*	Pyramid level 1*	4	±2.0SD	2x30+1 = 61	≈0.36mm	10	SSM
	Pyramid level 2*	4	±2.0SD	2x12+1 = 25	≈0.18mm	20	SSM
(2) Individual MM-LM ASM-fitting	MM	35	±1.0SD	2x8+1=16	≈0.18mm	25	SSM
	LM	35	±3.0SD	2x8+1=16	≈0.18mm	120	SSM
(3) Individual MM-LM ASM-relaxation	MM	NA**	NA**	2x20+1=41	≈0.18mm	10	Smoothing (15 it)
	LM	NA**	NA**	2x12+1=25	≈0.18mm	10	Smoothing (15 it)

* Performed using a 2 level Gaussian image pyramid (Level 1:0.729x0.729x1.397mm; Level 2: 0.365x0.365x0.698mm)

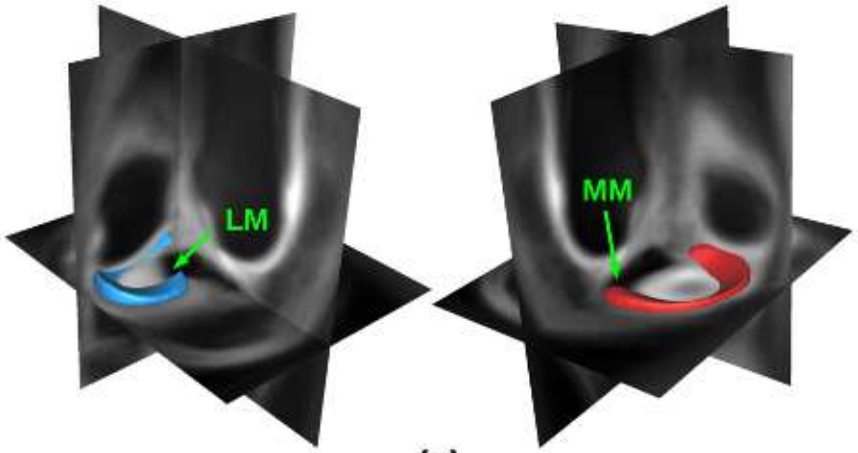
**No SSM utilised for the relaxation


MM
LM
Variation
1.000
0.500
0.000
Combined menisci
Medial meniscus
Lateral meniscus
$-3\sqrt{\lambda_k}$
0
$+3\sqrt{\lambda_k}$

(a)


Surface
Spacing s
P_0
P_1
P_N
Surface Normal
Inside Meniscus
+ Length l
- Length l

(b)


LM
MM

(c)

Supplementary Data A: Background method: training

From Dataset (A) at baseline, 85 *weDESS MR* examinations and associated manual MM and LM segmentations were selected to train our SSMs and image-feature models. Three patients exhibiting a destruction of more than 70% of the MM were excluded from the training to achieve higher statistical relevance. The method utilised to train the models was performed independently from the segmentation and involved 3 major stages: (1) statistical shape model training, (2) image-feature model training, and (3) affine average atlas image and surface training.

Statistical shape modelling

In this work SSMs of the MM, LM, and combined MM-LM were built from the initial dataset of manual segmentations following the method outlined by *Cootes et al.*[29]. MM and LM surfaces were reconstructed as 3D surfaces using the marching cube algorithm[47] and a set of *N*=85 menisci surfaces $M = \{M_0, \dots, M_{N-1}\}$ was obtained. In *M* each $M_i = \left(x_0^i, y_0^i, z_0^i, \dots, x_{ni}^i, y_{ni}^i, z_{ni}^i\right)$ represented a vector of n_i 3D points. Prior to statistical training, SSMs require point-wise correspondences to be established across all the surfaces. These were obtained by registering non-rigidly the MM and LM of M_0 onto all the other M_i using the Expectation Maximisation Iterative Closest Point algorithm[48] (EM-ICP). Each $S_i \in S$ was then expressed as a uni-dimensional vector of 3*n* components as defined in Eq. 3:

$$S_i = (x_0^i, y_0^i, z_0^i, \dots, x_n^i, y_n^i, z_n^i)^T \quad , \quad i = 0, \dots, N-1 \tag{3}$$

in which $\mathrm{k} = \left(\mathrm{x_k^i}, \mathrm{y_k^i}, \mathrm{z_k^i}\right)$ were the coordinates of the k^{th} point on the surface S_i and *n* the number of points in all the training surfaces. Absolute correspondences across S allowed the optimal shape alignment via Procrustes analysis[49], and the definition of a point distribution model (PDM). The mean shape $\bar{S}$ and the *3nx3n* covariance matrix C of the training set were computed from the PDM using Eq. (4):

$$\bar{S} = \frac{1}{N}\sum_{i=0}^{i<N} S_i \qquad ,and\,, \quad C = \frac{1}{N}\sum_{i=0}^{i<N}(S_i - \bar{S})(S_i - \overline{S})^T \tag{4}$$

The eigenvectors $p_1 = (k = 1, \dots, 3n)$ and eigenvalues $\lambda_k \in \lambda$ of *C* described the direction and the magnitude of the menisci shape variability across the atlas. Selecting the *t* largest λ_k allowed to model the most meaningful variations of the menisci while discarding variations associated to noise. Using standard principal component analysis[50], each S_i was then described as a weighted sum of $\bar{S}$ and the *t* major eigenvectors, as expressed in Eq. 5.

$$S_i^* = \bar{S} + Pb \tag{5}$$

in which $P = (p_1 \dots p_t)$ was the matrix of the *t* major eigenvectors, and $b = (b_1 \dots b_t)^T$ a vector of shape parameters. Hence, varying the values of b in an acceptable range allowed to generate likely meniscus shapes in an allowable shape domain. In this work, SSMs were further optimised by repeating the EM-ICP registration process using $\bar{S}$ instead of M_0 as initial registration surface.

As illustrated in Fig. S1a, the combined menisci SSM described the positional variability of the MM and LM, and individual MM and LM SSMs characterised the local variability.

[Suggestion supplementary Figure S1]

Image-feature model

The image-feature models used to drive the deformation of the SSMs were composed of the tissue intensity profiles surrounding the menisci in the training-set[30,31]. They provided priors on the intensity profiles likely to be found at each point of the menisci, and were generated from the pre-processed MRI using the surfaces S_i. For each S_i and each point $k = \left(x_k^i, y_k^i, z_k^i\right) \in S_i$, a one dimensional intensity profile $P_{i,k}$ of length *L* and spacing *s* was extracted along the surface normal in the positive and negative direction, and saved in the model. Separate image-feature models were generated for the MM, LM combined MM-LM, each containing *nx85*

likely menisci profiles of length *2L+1* (corresponding to the PDM). An illustration of grey level profiles extracted for *N* points of a given surface is provided on Fig. S1.b.

Affine average atlas image and surface

Average atlases have been shown to increase generalisability and accuracy of registration schemes. In this work, affine average atlas image and surface accounting for the population pose variability and morphology were generated and used to robustly initialise the ASM-fitting stage. They were obtained by registering affinely the pre-processed images to a common image and averaging the results into a 'knee average image. The affine transformations obtained were then propagated to the respective surfaces S_i and the mean shape (MM and LM) was calculated in the atlas space using Eq. 4. The average atlas image and surface are illustrated in Fig. S1.c.