

Development of Image Overlay and Knowledge Transfer Module Technologies Aimed at Enhancing Feasibility and External Validation of Magnetic Resonance Imaging-based Scoring Systems

Jacob L. Jaremko, Meaghan Pitts, Walter P. Maksymowych, and Robert G. Lambert

ABSTRACT. Objective. Semiquantitative arthritis scoring assesses disease burden by scoring presence/extent of features such as bone marrow lesion (BML) or effusion in multiple anatomic regions at a joint. An image overlay clarifying region borders may enhance feasibility and reliability of these scoring systems. To be scalable for use in large clinical trials, systematic computer-based user training is desirable. We developed an overlay and user training module for magnetic resonance imaging (MRI)-based scoring of hip osteoarthritis (OA).

Methods. We designed a semitransparent 2-dimensional image overlay applied to individual MRI slices to facilitate hip OA scoring [HIMRISS (Hip Inflammation MRI Scoring System)], initially using freeware and then in a customized HTML Web browser environment. We developed a systematic knowledge translation package including instructional presentation, fully scored expert consensus cases, and video tutorials for training in the use of these scoring systems with the overlays. Three musculoskeletal radiologists who had not used this scoring system before each performed a scoring exercise with no overlay, then repeated this with overlays after completing the training module. Based on postexercise interviews and a reader survey, we identified and corrected problems in the module. The entire training process was then repeated using 3 new readers.

Results. Overlays were considered useful, particularly when integrated into a Web browser. The knowledge translation module was considered conceptually valuable, but as initially implemented was too lengthy and not sufficiently interactive.

Conclusion. Semitransparent image overlays and standardized knowledge translation modules for reader training show promise to facilitate reader calibration using MRI-based scoring systems. Based on our experience, knowledge translation modules should emphasize close feedback evaluating performance and reader time efficiency. (J Rheumatol First Release July 1 2015; doi:10.3899/jrheum.141573)

Key Indexing Terms:

OSTEOARTHRITIS
DIAGNOSIS

MAGNETIC RESONANCE IMAGING

HIP
VALIDATION

As the range of therapeutic options for arthritis increases, there is a pressing need for objective outcome measures that can show changes in the time frame of clinical trials. Magnetic resonance imaging (MRI)-based semiquantitative arthritis scoring systems fulfill this need by assessing disease

burden through scoring presence and/or extent of features such as bone marrow lesions (BML) or effusion in multiple anatomic regions within a joint. There has recently been a proliferation of such semiquantitative scoring systems. For example, at the hand and wrist, the OMERACT Rheumatoid Arthritis MRI Score (RAMRIS) is available to score rheumatoid arthritis (RA) severity^{1,2} and has been applied in multiple randomized controlled trials, with highly reliable and sensitive results encouraging its adoption as a formal outcome measure in RA³. RAMRIS has also been applied to other joints such as the feet⁴, and to other arthropathies such as gout⁵. A similar semiquantitative scoring approach has recently been applied to osteoarthritis (OA). At the knee, the Whole Organ MRI Score⁶, Boston Leeds OA Score⁷, and MOAKS (MRI OA Knee Score⁸) systems represent successive iterations of a comprehensive whole-organ scoring system, while KIMRISS (Knee Inflammation MRI Scoring

From the Department of Radiology and Diagnostic Imaging and the Department of Rheumatology, Faculty of Medicine, University of Alberta, Edmonton, Alberta, Canada.

Supported by a Capital Health Endowed Chair in Diagnostic Imaging.

J.L. Jaremko, FRCPC; M. Pitts, FRCPC, Department of Radiology and Diagnostic Imaging, Faculty of Medicine, University of Alberta; W.P. Maksymowych, FRCPC, Department of Rheumatology, Faculty of Medicine, University of Alberta; R.G. Lambert, FRCPC, Department of Radiology and Diagnostic Imaging, Faculty of Medicine, University of Alberta.

Address correspondence to Dr. J.L. Jaremko, Department of Radiology and Diagnostic Imaging, 2A2.41 WC Mackenzie Health Sciences Centre, University of Alberta, 8440-112 St., Edmonton, Alberta T6G 2B7, Canada. E-mail: jjaremko@ualberta.ca

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2015. All rights reserved.

System⁹) focuses on potentially treatable active pathology, which may include a component of inflammation. For hip OA, whole-organ scoring by HOAMS (Hip OA MRI Scoring System)¹⁰, or focused inflammatory scoring by HIMRISS (Hip Inflammation MRI Scoring System^{11,12}) are available. The SpondyloArthritis Research Consortium of Canada score is available for inflammatory arthritis involving sacroiliac joints¹³ and spine¹⁴.

The above systems all evaluate BML, which is an MRI finding in periarticular bone closely correlated with symptoms¹⁵ and prognosis^{16,17} in OA. Each system divides bone into regions for scoring. In RAMRIS the region borders are natural: each small bone of the hand and wrist is scored separately¹. Larger bones at the hip, knee, or sacroiliac joints are subdivided into visually manageable regions. In some systems (HIMRISS, KIMRISS) the reader makes binary decisions whether BML is present or absent in many small regions^{9,11}, while in others (e.g., HOAMS, MOAKS) the reader counts foci of BML in each of a few larger regions and the percentage of each region containing BML^{8,10}. Each scoring system has an associated diagram demonstrating the relevant region borders (Figure 1). Many of these decisions could potentially be simplified and made more reliable by having a semitransparent grid overlay placed directly on the digital imaging and communications in medicine (DICOM) (digital imaging and communications in medicine) images, more clearly and reproducibly delineating boundaries for each region to be scored.

Systematic user training has been shown to improve intra- and inter-reader reliability for semiquantitative scoring systems, whether based on radiographs¹⁸ or MRI¹⁹. For RAMRIS, the most established MRI-based system, extensive training materials exist that are primarily based on a published comprehensive reference image atlas^{20,21,22}. Two readers improved their MRI erosion score intraclass correlation coefficients from 0.52–0.60 to 0.77–0.89 after training provided directly by an experienced reader, consisting of a 3-h initial session explaining anatomy, pathophysiology, and scoring strategies, and two 2-h calibration sessions reviewing real cases in consensus¹⁹. User training is not as universally rigorous as this in published studies²³. For HOAMS, 2 expert readers performed two 2-h calibration sessions against each other prior to a validation study, achieving κ 0.8510, while for MOAKS, 2 readers performed a single calibration session on 10 cases, achieving κ 0.72–1.0 for BML⁸. There is little published literature on the effectiveness of these scoring systems in novice or expert readers, or on how best to train and calibrate readers. For large-scale use in clinical trials and other research, it is desirable for reader training to be computer-based to allow readers to train at their leisure and to minimize time demands on the original developers.

In this report we describe the development of 2 innovations in MRI-based semiquantitative arthritis scoring aimed at enhancing reader reliability: semitransparent image over-

lays and a computer-based knowledge translation module. The innovations are applied to HIMRISS, a scoring system for hip OA, but conceptually can be applied to any MRI-based scoring systems existing or under development. This work was prepared for the MRI in hip OA special interest group, at the Outcome Measures in Rheumatology (OMERACT) 12 conference (Budapest, May 7–11, 2014). A second related article describing a multireader validation exercise specifically assessing reliability of HIMRISS, without and with overlay, and knowledge transfer module is published elsewhere in these proceedings²⁴.

MATERIALS AND METHODS

HIMRISS

An introduction to the HIMRISS has been published^{11,12}. HIMRISS grades BML and synovitis/effusion, 2 MRI features thought to reflect potentially treatable active pathology (which may or may not include a frank inflammatory component¹¹) in hip OA. HIMRISS requires a coronal fluid-sensitive MRI sequence [typically short-tau inversion recovery (STIR)] of 3–4 mm thickness, and a matched coronal T1-weighted sequence to help assess bone anatomy and lesions. Each hip is scored separately. First, the central slice of the hip is determined, as the middle slice between the most anterior and most posterior slices on which femoral head marrow is visible; if the mathematical middle is not a whole number, it is rounded to pick the slice just anterior to the middle. Each region of the femoral head and acetabulum on each of 15 slices (7 anterior and 7 posterior to the central slice) is assigned a binary score of 0 (no BML) or 1 (BML present). Regions are designed to be about similar in size, so that on the central 5 slices, 9 femoral and 3 acetabular regions are scored while on the anterior and posterior slices (up to 5 each), 2 femoral and 2 acetabular regions are scored, for a total score of up to 100 (Figure 2). A score of 1 is given to each region containing increased fluid signal within bone marrow that does not represent red marrow or cyst (Figure 3). If an overlay is used, it is placed to best fit the femoral head on the central slice, and left in the same position when reviewing all the other slices.

Effusion/synovitis is scored on each of the 15 slices, by measuring the maximum depth of STIR-intense (i.e., bright) fluid/synovial membrane surrounding the hip. The short axis of the largest fluid recess seen on each slice is measured. Each slice is scored 0, 1, or 2 for thickness < 2 mm, 2–3.9 mm, or \geq 4 mm, respectively, for a total possible score of up to 30 (Figure 4).

Overlay Technology

The HIMRISS overlay is a single image combining templates for the femoral head and acetabulum. Its lines and curves, with no distracting background, are visually superimposed over the DICOM MR images being viewed by the reader, as if a physical transparency was placed over the computer screen. Our initial feasibility studies used freeware for Windows (Vitrite, available from: www.vanmiddlesworth.org/vitrite/) and Macintosh (Overlay2, available from: www.colinthomas.com/overlay/), which allow placement of a window “always on top” over other windows, and permit rescaling and multilevel adjustment of window transparency. The user opens the DICOM viewer, then opens the overlay image in another window, resizes that window to “fit” the displayed hip, and makes the image semitransparent, then reads each image slice (Figure 5). As a subsequent platform-independent refinement, we integrated code into an HTML environment (available from: <http://carearthritits.com>; Figure 6) containing an image viewer based on open-source elements from ClearCanvas (v. 2.0, Cleome). The freeware programs require several mouse clicks to set up the overlay, but have the advantage of allowing use of any DICOM viewer. The HTML version requires use of the integrated DICOM viewer, but is automatically set up, with the only required user action being to move and resize the overlay to best fit the joint.

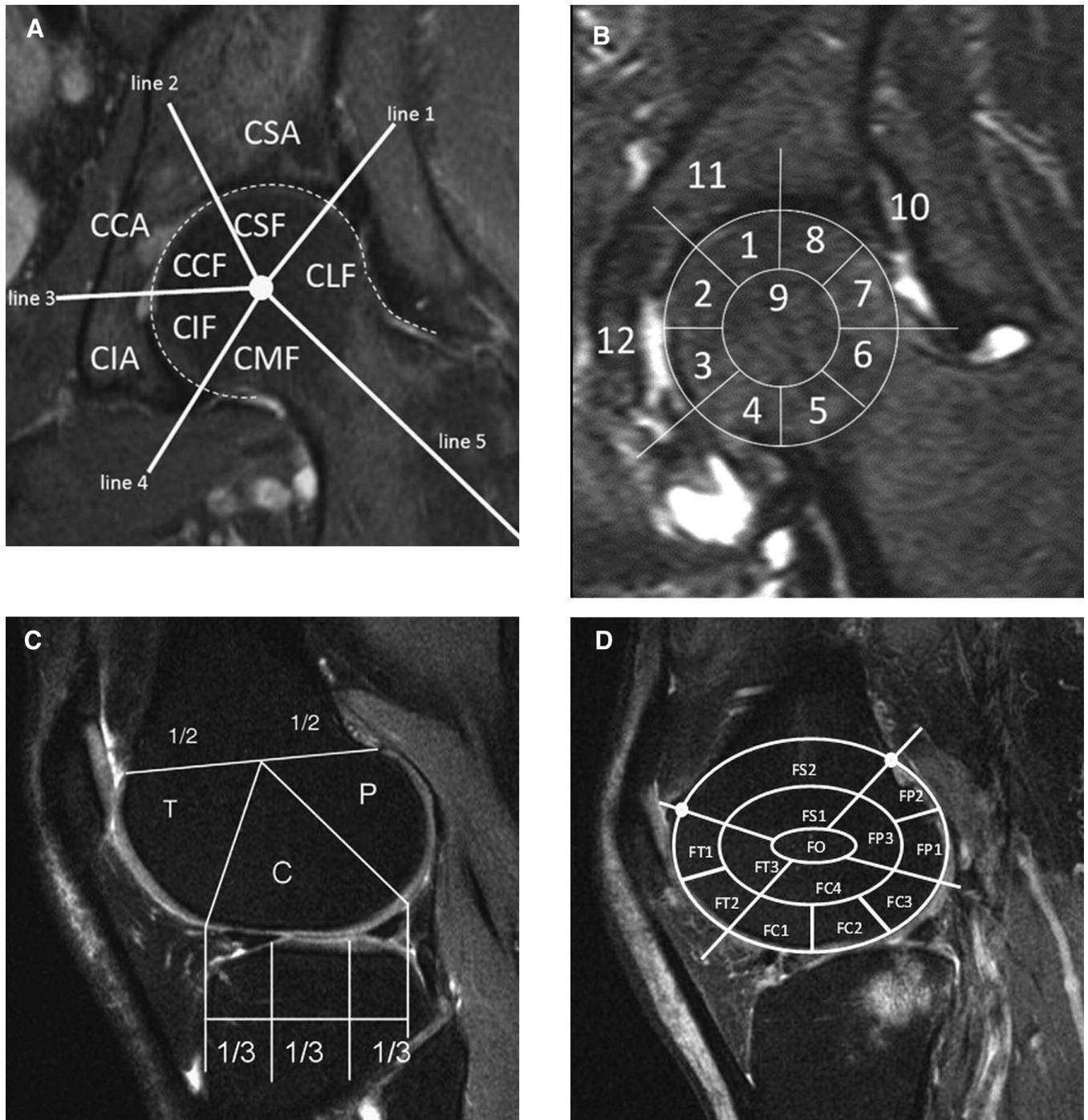


Figure 1. Diagrams demonstrating the diversity of regional subdivisions for magnetic resonance imaging (MRI)-based semiquantitative arthritis scoring systems. A. HOAMS at the hip¹⁰; B. HIMRISS at the hip¹¹; C. MOAKS at the knee⁸; D. KIMRISS at the knee (under development). Although the specific regional divisions and other scoring rules differ between each system, each shares the principle of assigning each region a value for the imaging features contained therein, especially the presence and extent of bone marrow lesions. HOAMS: Hip Osteoarthritis Magnetic Resonance Image Scoring System; HIMRISS: Hip Inflammation MRI Scoring System; MOAKS: MRI OA Knee Score; KIMRISS: Knee Inflammation MRI Scoring System; OA: osteoarthritis.

MRI and Clinical Data Sources

Data were extracted from patients selected randomly from an ongoing prospective study in which hip MRI, clinical history, and examination are obtained pre- and 8 weeks post-intraarticular steroid injection in patients with hip OA meeting American College of Rheumatology criteria²⁵. Clinical

data included Western Ontario and McMaster Universities Arthritis Index scores for pain, stiffness, and function^{26,27} at each visit. Coronal STIR MR images were available with slice thickness 4 mm, field of view 400 × 400 mm, matrix size 512–256, TR 3550, TE 51, TI 145 ms. Both hips were included for each patient.

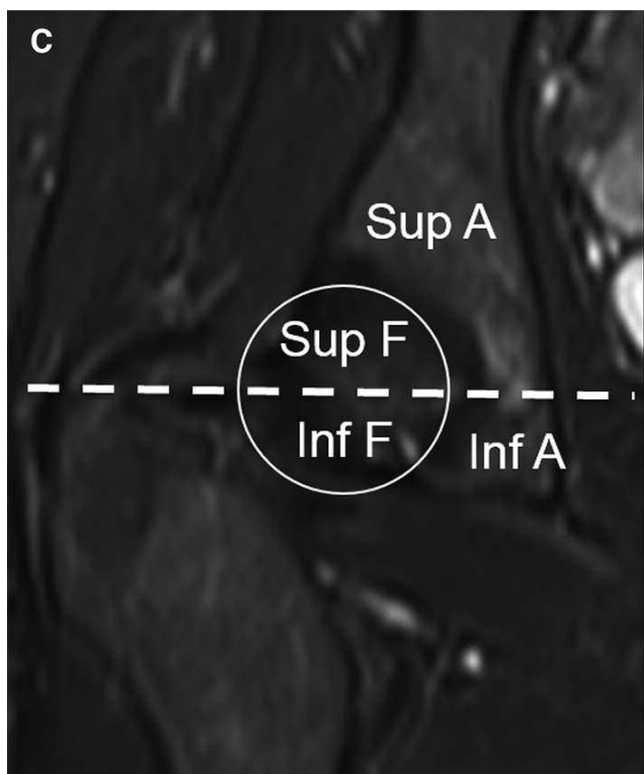
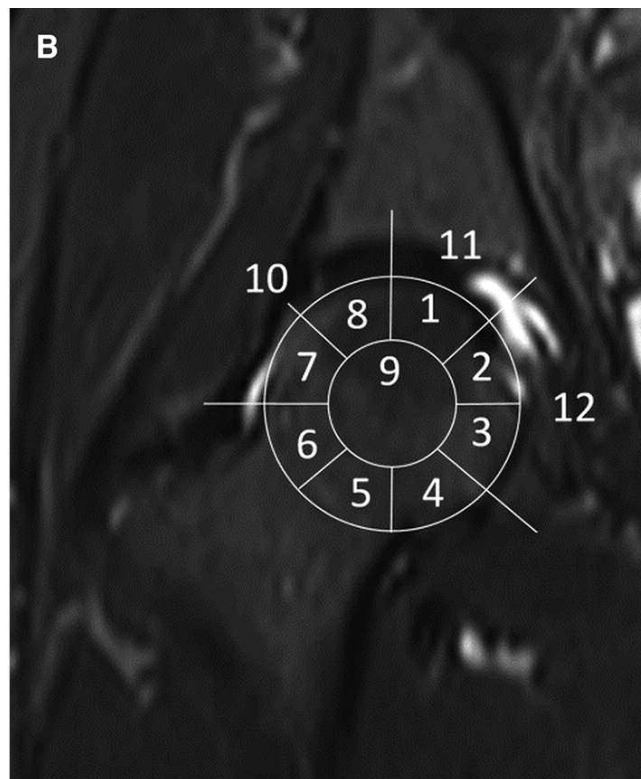
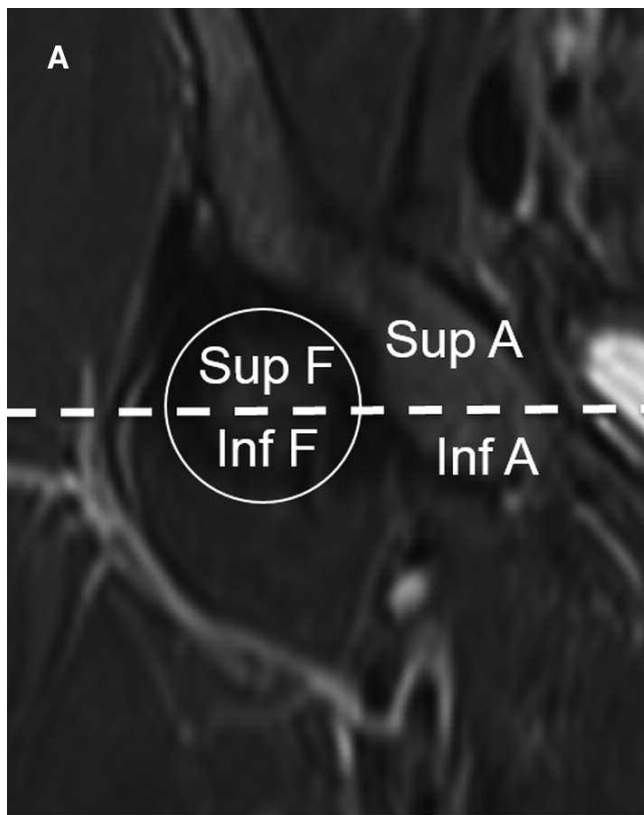


Figure 2. HIMRISS regions for scoring BML. A. Anterior 5 slices. B. Central 5 slices. C. Posterior 5 slices. For example, if the central slice is #9, the slices to be scored are 2–6 anterior, 7–11 central, 12–16 posterior. The regions to be scored are fewer and simpler on the anterior and posterior slices than the central slices. If no bone is present in a region on a given slice, or the slice does not exist, the region(s) are scored 0. HIMRISS: Hip Inflammation Magnetic Resonance Image Scoring System; BML: bone marrow lesion.

Expert Consensus Exercise

Two expert readers from the team developing HIMRISS (RL and JJ, both musculoskeletal radiologists with 30 and 10 years imaging experience,

respectively) scored both hips at initial and followup visits for 8 patients (i.e., 32 hips), and performed a detailed consensus review of scores for each grid cell of each slice of each scan, resolving discrepancies by consensus

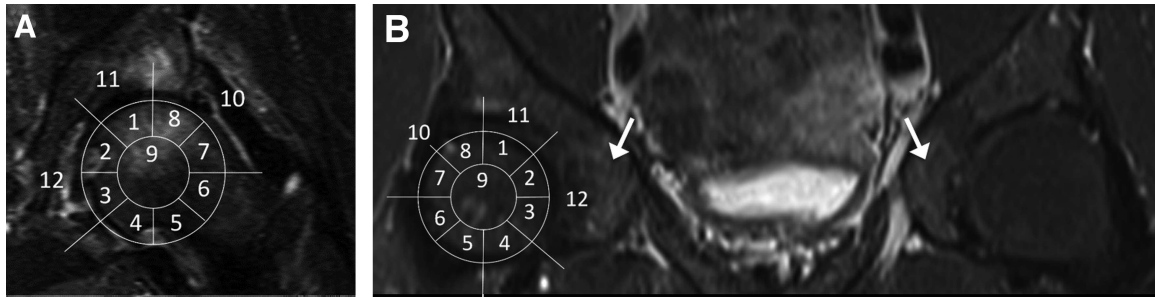


Figure 3. HIMRISS BML scoring examples. A. A score of 1 is given to each region containing any fluid signal within bone marrow greater than expected for red marrow, and not representing simply a cyst. Here, regions 1, 7, 8, and 9 are scored 1. Regions 2 and 3 contain high signal medially, but this is the fovea centralis and not BML. Regions 4-5 contain high signal inferiorly but this is outside bone, in the inferior articular recess, and not BML. Regions 10 and 11 of the acetabulum are also scored 1. B. BML is present in the femoral head sectors 8 and 9, and acetabular sectors 10 and 11. Signal is mildly bright in sector 12 near the arrow, but similar to the other hip (arrow), and in distribution typical of red marrow, commonly seen in sector 12 (as well as sectors 4 and 5). This is scored 0. HIMRISS: Hip Inflammation Magnetic Resonance Image Scoring System; BML: bone marrow lesion.

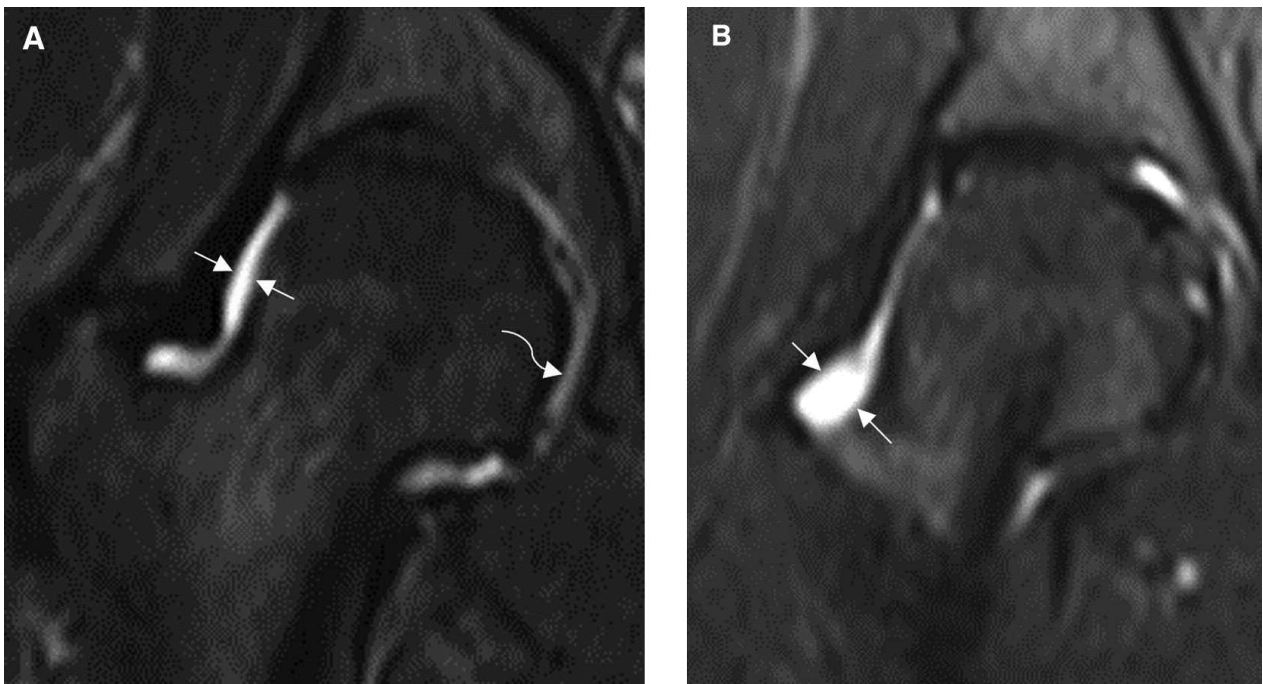


Figure 4. Effusion scoring on HIMRISS. A. The short axis of this high-intensity area, representing a combination of fluid and synovium, measures 2.6 mm for a slice score of 1. Note that the moderately T2-bright area at the curved arrow is not fluid or synovium, but cartilage, and is not scored. B. This wider recess (arrows) measures 8 mm, for a slice score of 2. HIMRISS: Hip Inflammation Magnetic Resonance Image Scoring System.

and providing comments for difficult cells. Initial and followup scans were reviewed side by side, using randomized timepoints so the readers were unaware which study came first.

Knowledge Translation Module Development

We developed a knowledge translation module to train new readers. This module comprised the following:

1. An 85-slide PowerPoint presentation introducing the rationale for scoring BML and effusion/synovitis, the method of determining the central slice and placing the overlay template, examples of scoring BML and synovitis/effusion including classic and borderline cases and common pitfalls, sample scoring sheets, and instructions on overlay use. This was developed by a musculoskeletal radiology OMERACT fellow (MP) over a 6-month period

with frequent review by study investigators to ensure accuracy of the information presented.

2. Fully scored consensus cases in DICOM format (8 patients \times 2 hips \times 2 timepoints = 32 hips), a blank scoring spreadsheet, and fully scored expert scoring spreadsheet in Microsoft Excel 2013 format.

3. A 5-minute video tutorial (suitable for Internet download or podcast) demonstrating the overlay and scoring system.

Our intent was for readers to (1) review the PowerPoint presentation and instructional video; (2) score 1-2 of the provided cases; (3) compare scores to expert consensus and evaluate discrepancies to bring scores closer to expert scores; (3) repeat on the next 1-2 cases; and (4) repeat further in case of continued substantial discrepancies.

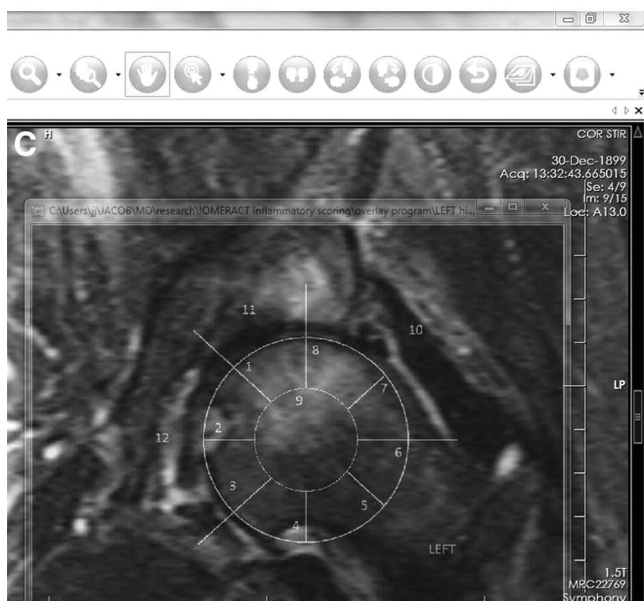
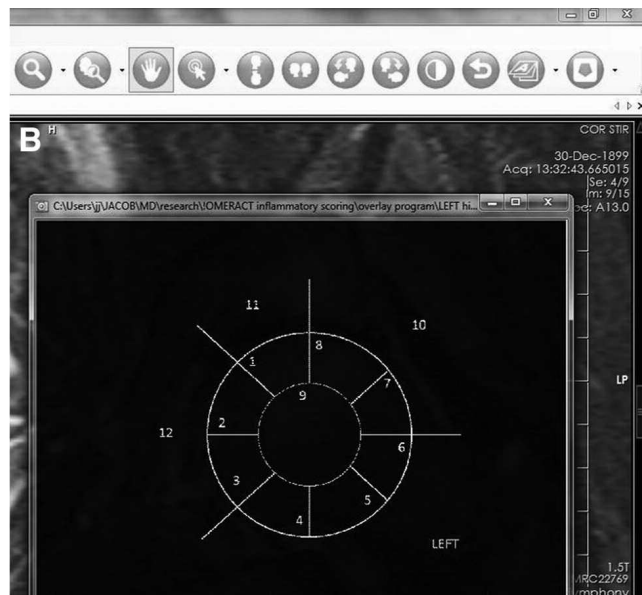
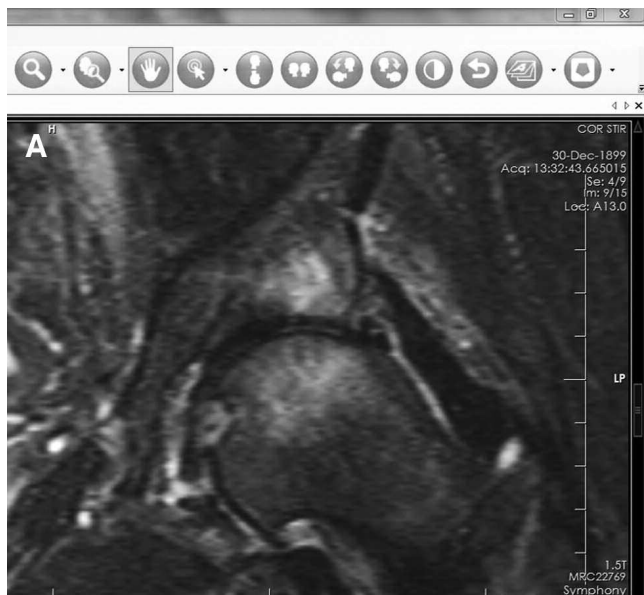


Figure 5. Use of an image overlay: freeware method. A. The reader opens the MRI in the DICOM viewer of choice (here, ClearCanvas v. 2.0), with the image overlay program running in the background (not visible). B. The overlay image is opened and placed over the MRI by moving its window. C. The image overlay program is activated to make that window “always on top” and nearly transparent. The reader makes fine adjustments to the window position and size as required, then changes focus to the DICOM viewer and scrolls through all images, performing scoring. MRI: magnetic resonance imaging; DICOM: digital imaging and communications in medicine.

Reading Exercise Design

We recruited 3 musculoskeletal radiologists (NW from UK, MP from Australia, DM from Canada, with 8, 10, and 6 years experience, respectively) who had not previously used the HIMRISS scoring system. They were given the published manuscript¹¹ describing HIMRISS, and without using the overlay or learning module, they scored both hips, scanned pre- and post-steroid injection of 1 hip, in 16 patients. After 2 weeks to review the knowledge translation module, they repeated the scoring exercise with an additional 8 cases intermixed to avoid recollection bias. They completed a postsurvey questionnaire regarding the scoring system, exercise and knowledge translation module, and postexercise interviews were conducted. The 2 expert readers who prepared the consensus cases were also interviewed.

RESULTS

Image Overlays

Overlays were considered straightforward to use and helpful

(“good” to “excellent”) by 4/5 readers. The fifth reader felt the overlay was more useful as a conceptual guide, rather than to physically place it over the images, and referred to a paper printout of the overlay while scoring cases. Readers preferred the HTML version to the use of freeware programs because of the convenience of having the overlay program automatically integrated into HTML. Although there were limited features available for image analysis on the HTML-integrated DICOM viewer, no reader commented that they felt restricted by this.

Knowledge Translation Module

The readers were all full-time radiologists with busy clinical practices. Readers spent 15–60 min reviewing the intro-

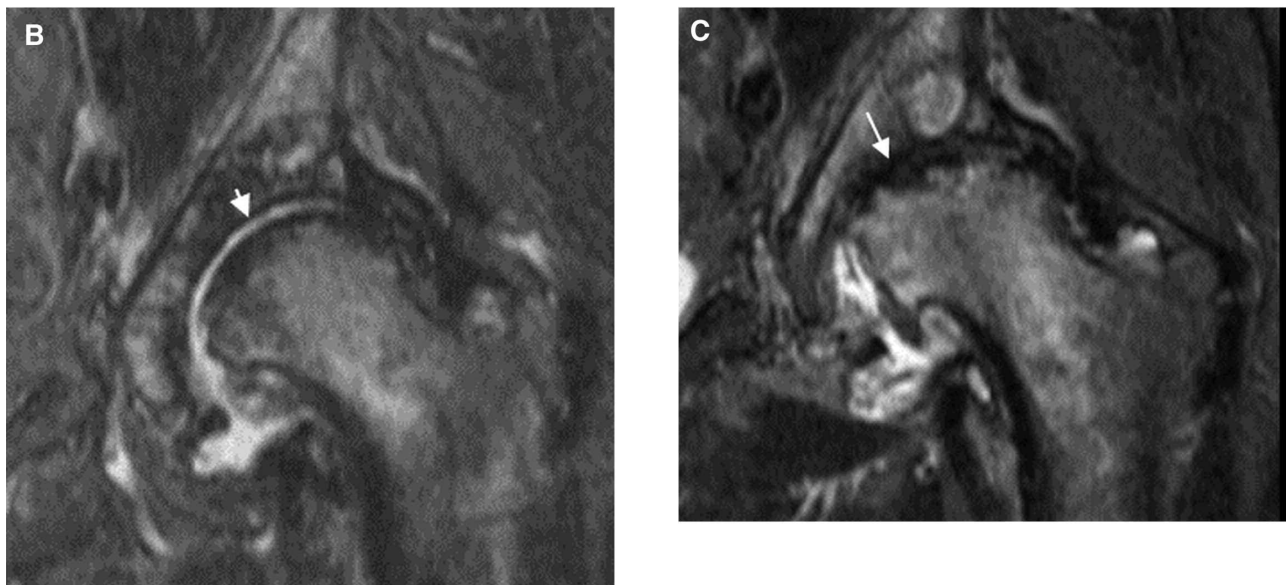
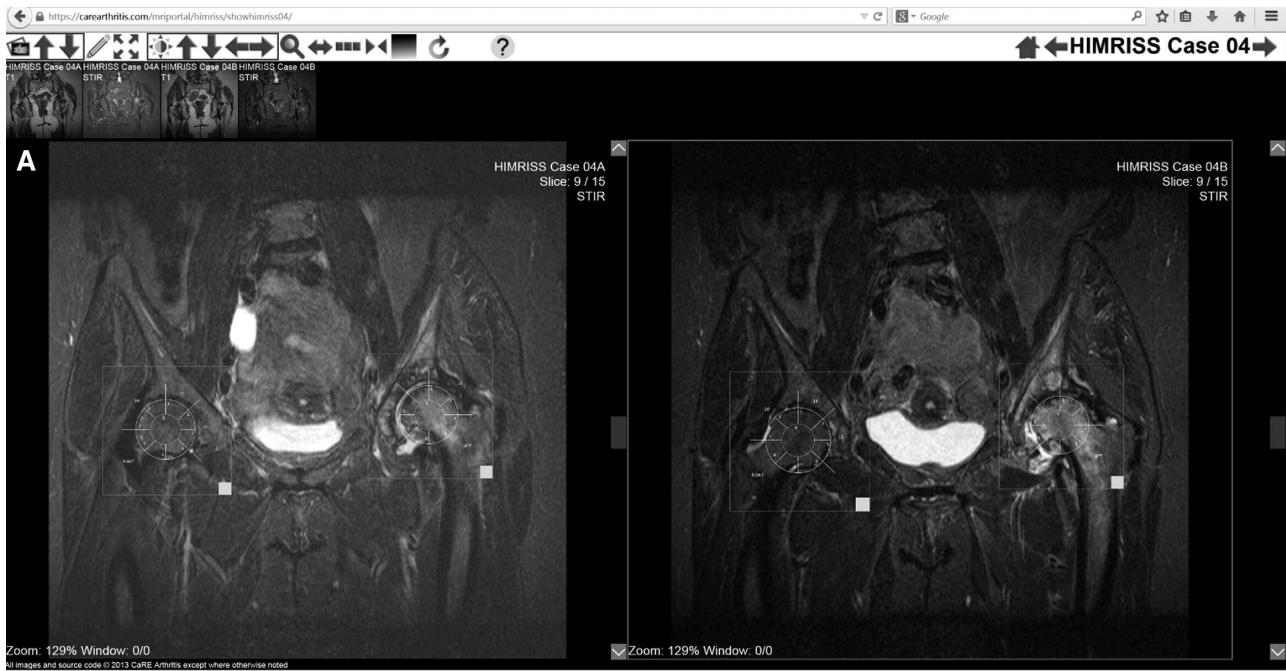


Figure 6. Use of an image overlay: A. HTML environment as implemented on the Website www.carearthris.com. Image series from the 2 timepoints are placed side by side and scroll together for direct comparison of interval change. The overlays are placed at both hips in each series by dragging and resizing them with mouse clicks on the small white squares at their lower right corners. B–C. Note that in this case there is severe left hip osteoarthritis with obvious adverse change from B to C, showing interval progressive collapse of the femoral head and loss of the joint line, clearly visible in B (arrow) and obliterated in C (arrow). There are extensive bone marrow lesions and joint effusion at both timepoints.

ductory presentation. Comprehensibility was rated “good” to “excellent,” and usefulness was rated “fair” to “excellent.” Readers commented favorably on inclusion of examples, and specific learning issues were corrected; for example, before reviewing this presentation 1 reader had understood effusion scoring to be binary when in fact scores of 0, 1, or 2 are possible for each slice. One reader noted that while overlay

placement and BML scoring were “self-explanatory,” effusion scoring was more difficult to understand.

Time spent by readers reviewing consensus cases was 10–120 min. Two readers reviewed 2 cases at 2 timepoints, and 1 reader reviewed a single hip from all 8 cases at 1 timepoint. Usefulness of the cases was rated “good” to “excellent.” Readers felt that too many cases were presented

and there was not enough time to review them all. One reader noted that he had difficulty achieving consensus with the expert scores for effusions even after reviewing the consensus cases.

Considering the entire training module, readers agreed or strongly agreed that review of the module was essential for reliable scoring. All 3 readers commented that the time frame for review (2 weeks to perform image readings, 2 weeks for learning module, 2 weeks for new readings) was challenging with their busy schedules.

At the OMERACT 12 hip OA special interest group meeting, the overlays and knowledge translation module were presented to participants. Group discussion centered on a perceived need to develop more systematic user training, and resulted in consensus to raise a final voting question to the entire OMERACT group: "Should a systematic approach to knowledge translation and user training be applied to MRI-based scoring systems?" This received a positive vote of 65%.

Based on feedback from this process, we made multiple changes to training. A key improvement was an interactive spreadsheet in which readers fill in scores for each consensus case, then receive instant feedback through color-coding of cells highlighting the discrepancies between their score and expert consensus. The reader continues taking new cases until total score varies by less than 10% from expert consensus. We also revised the training presentation to clarify effusion scoring, and added a short series of multiple-choice comprehension questions at the end of the presentation.

DISCUSSION

We presented and evaluated the feasibility and potential added value of 2 innovations for MRI-based arthritis scoring: an image overlay and a knowledge translation module for training.

The image overlay was straightforward to implement and was reviewed favorably by expert and naive readers regarding feasibility and potential value. This justifies us proceeding to more formally evaluate the effect of the overlay on reader reliability in future. We also learned that readers prefer an integrated image review/overlay environment, even if the image viewer is limited in functionality. This will direct further overlay development and extension to other scoring systems.

The knowledge translation module was also considered useful by readers, but we learned there were important flaws in its implementation. The PowerPoint presentation was well received, and there was evidence that readers did learn from this. The points of confusion noted by readers in our explanation of effusion scoring were rectified. We were surprised how few consensus cases (4–8 of the available 32 fully scored hips) each reader reviewed. Likely contributing to this, the time given for module review (2 weeks) was short given our readers' busy schedules. Also, readers lacked motivation

to complete all consensus cases because they were not required to submit any scores for them, and methods often seemed "self-explanatory" after review of the presentation. Our new interactive spreadsheet will encourage more complete review of consensus cases, because it provides automatic cell-by-cell comparisons of each reader's scores with expert scores, and readers must submit their scores demonstrating achievement of a threshold of reliability. To further motivate readers to improve, reader scores are compared to expert scores with intraclass correlation coefficients provided immediately. This approach has been implemented for scoring of the modified Stoke Ankylosing Spondylitis Spine score (available from: www.carearthritis.com/Radiography.php)²⁸. Since completing the modified module now takes more time than spent by readers originally, ample time needs to be available for busy readers to fit this into their schedule, e.g., 4–6 weeks rather than the 2 weeks that we allowed.

Although the image overlay and knowledge translation module presented here were prepared for the HIMRISS scoring system, the principles behind each of these innovations can be applied to any MRI-based scoring system. This has potential to improve and standardize scoring reliability, which is vital as MRI assumes an increasingly important role as an endpoint for prognostic studies and clinical trials.

Semitransparent image overlays and standardized knowledge translation modules for reader training are 2 innovations showing promise to facilitate MRI scoring of arthritis at the hip and knee. Future development of image overlays can focus on integrating these within a unified Web-based image viewing environment. Refinements to knowledge translation modules for reader training should focus on iterative feedback linking reader performance to expert scores as tightly as possible, and on providing appropriate time and guidance to complete each task.

ACKNOWLEDGMENT

We thank Joanne McGoey, Elizabeth Fehlauer, Amy Norton, and Linda Woodhouse for their hard work and valuable contributions to this study; Naomi Winn, Consultant Radiologist, University of Manchester, UK; Marcus Pianta, Consultant Radiologist, St. Vincent's Hospital, Victoria, Australia; David McDougall, Department of Radiology, University of Alberta, Canada, for contributing as readers; and Joel Paschke of CaRE Arthritis Ltd. for computer programming.

REFERENCES

1. Ostergaard M, Peterfy C, Conaghan P, McQueen F, Bird P, Ejbjerg B, et al. OMERACT rheumatoid arthritis magnetic resonance imaging studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. *J Rheumatol* 2003;30:1385-6.
2. Haavardsholm EA, Ostergaard M, Ejbjerg BJ, Kvan NP, Uhlig TA, Lilleas FG, et al. Reliability and sensitivity to change of the OMERACT rheumatoid arthritis magnetic resonance imaging score in a multireader, longitudinal setting. *Arthritis Rheum* 2005;52:3860-7.

3. Peterfy C, Ostergaard M, Conaghan PG. MRI comes of age in RA clinical trials. *Ann Rheum Dis* 2013;72:794-6.
4. Baan H, Bezooijen R, Avenarius JK, Dubbeldam R, Drossaers-Bakker WK, van de Laar MA. Magnetic resonance imaging of the rheumatic foot according to the RAMRIS system is reliable. *J Rheumatol* 2011;38:1003-8.
5. Popovich I, Dalbeth N, Doyle A, Reeves Q, McQueen FM. Exploring cartilage damage in gout using 3-T MRI: distribution and associations with joint inflammation and tophus deposition. *Skeletal Radiol* 2014;43:917-24.
6. Peterfy CG, Guermazi A, Zaim S, Tirman PF, Miaux Y, White D, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. *Osteoarthritis Cartilage* 2004;12:177-90.
7. Hunter DJ, Lo GH, Gale D, Grainger AJ, Guermazi A, Conaghan PG. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). *Ann Rheum Dis* 2008;67:206-11.
8. Hunter DJ, Guermazi A, Lo GH, Grainger AJ, Conaghan PG, Boudreau RM, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). *Osteoarthritis Cartilage* 2011;19:990-1002.
9. Maksymowych W, Weber U, Pianta M, Lambert R. Semi-quantitative assessment of bone marrow edema and synovitis-effusion in osteoarthritis with the knee inflammation MRI scoring system (KIMRISS): A target lesion based methodology [abstract]. *Ann Rheum Dis* 2012;71 Suppl 3:585.
10. Roemer FW, Hunter DJ, Winterstein A, Li L, Kim YJ, Cibere J, et al. Hip Osteoarthritis MRI Scoring System (HOAMS): reliability and associations with radiographic and clinical findings. *Osteoarthritis Cartilage* 2011;19:946-62.
11. Jaremko JL, Lambert RG, Zubler V, Weber U, Loeuille D, Roemer FW, et al. Methodologies for semiquantitative evaluation of hip osteoarthritis by magnetic resonance imaging: approaches based on the whole organ and focused on active lesions. *J Rheumatol* 2014;41:359-69.
12. Maksymowych WP, Cibere J, Loeuille D, Weber U, Zubler V, Roemer FW, et al. Preliminary validation of 2 magnetic resonance image scoring systems for osteoarthritis of the hip according to the OMERACT filter. *J Rheumatol* 2014;41:370-8.
13. Landewe RB, Hermann KG, van der Heijde DM, Baraliakos X, Jurik AG, Lambert RG, et al. Scoring sacroiliac joints by magnetic resonance imaging. A multiple-reader reliability experiment. *J Rheumatol* 2005;32:2050-5.
14. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Ostergaard M, et al. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. *J Rheumatol* 2007;34:862-70.
15. Taljanovic MS, Graham AR, Benjamin JB, Gmitro AF, Krupinski EA, Schwartz SA, et al. Bone marrow edema pattern in advanced hip osteoarthritis: quantitative assessment with magnetic resonance imaging and correlation with clinical examination, radiographic findings, and histopathology. *Skeletal Radiol* 2008;37:423-31.
16. Davies-Tuck ML, Wluka AE, Forbes A, Wang Y, English DR, Giles GG, et al. Development of bone marrow lesions is associated with adverse effects on knee cartilage while resolution is associated with improvement—a potential target for prevention of knee osteoarthritis: a longitudinal study. *Arthritis Res Ther* 2010;12:R10.
17. Tanamas SK, Wluka AE, Pelletier JP, Pelletier JM, Abram F, Berry PA, et al. Bone marrow lesions in people with knee osteoarthritis predict progression of disease and joint replacement: a longitudinal study. *Rheumatology* 2010;49:2413-9.
18. van der Heijde D, Boers M, Lassere M. Methodological issues in radiographic scoring methods in rheumatoid arthritis. *J Rheumatol* 1999;26:726-30.
19. Bird P, Joshua F, Lassere M, Shnier R, Edmonds J. Training and calibration improve inter-reader reliability of joint damage assessment using magnetic resonance image scoring and computerized erosion volume measurement. *J Rheumatol* 2005;32:1452-8.
20. Ostergaard M, Edmonds J, McQueen F, Peterfy C, Lassere M, Ejbjerg B, et al. An introduction to the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. *Ann Rheum Dis* 2005;64 Suppl 1:3-7.
21. Ejbjerg B, McQueen F, Lassere M, Haavardsholm E, Conaghan P, O'Connor P, et al. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the wrist joint. *Ann Rheum Dis* 2005;64 Suppl 1:23-47.
22. Conaghan P, Bird P, Ejbjerg B, O'Connor P, Peterfy C, McQueen F, et al. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the metacarpophalangeal joints. *Ann Rheum Dis* 2005;64 Suppl 1:11-21.
23. Reichmann WM, Maillefert JF, Hunter DJ, Katz JN, Conaghan PG, Losina E. Responsiveness to change and reliability of measurement of radiographic joint space width in osteoarthritis of the knee: a systematic review. *Osteoarthritis Cartilage* 2011;19:550-6.
24. Maksymowych WP, Pitts M, Budak MJ, Gracey D, Lambert RG, McDougall D, et al. Development and preliminary validation of a digital overlay-based learning module for semiquantitative evaluation of magnetic resonance imaging lesions in osteoarthritis of the hip. *J Rheumatol* 2015;42: in press.
25. Altman R, Alarcon G, Appelrouth D, Bloch D, Borenstein D, Brandt K, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. *Arthritis Rheum* 1991;34:505-14.
26. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833-40.
27. Bellamy N. WOMAC: a 20-year experiential review of a patient-centered self-reported health status questionnaire. *J Rheumatol* 2002;29:2473-6.
28. Maksymowych WP, Learch T, Lambert RG, Ward M, Haroon N, Inman R, et al. Development and validation of the spondyloarthritis radiography module for calibration of readers using the modified Stoke Ankylosing Spondylitis Score. *Arthritis Care Res* 2014; 66:55-62.