Development and Validation of an OMERACT MRI Whole-Body Score for Inflammation in Peripheral Joints and Entheses in Inflammatory Arthritis (MRI-WIPE)

Simon Krabbe, Iris Eshed, Frédérique Gandjbakhch, Susanne J. Pedersen, Paul Bird, Ashish J. Mathew, Robert G. Lambert, Walter P. Maksymowych, Daniel Glinatsi, Maria S. Stoenoiu, René Poggenborg, Lennart Jans, Jacob L. Jaremko, Nele Herregods, Violaine Foltz, Philip G. Conaghan, Christian E. Althoff, Joel Paschke, Charles Peterfy, Kay-Geert A. Hermann, and Mikkel Østergaard, for the OMERACT MRI in Arthritis Working Group

ABSTRACT. Objective. To develop a whole-body magnetic resonance imaging (MRI) scoring system for peripheral arthritis and enthesitis.

Methods. After consensus on definitions/locations of MRI pathologies, 4 multireader exercises were performed. Eighty-three joints were scored 0–3 separately for synovitis and osteitis, and 33 entheses 0–3 separately for soft tissue inflammation and osteitis.

Results. In the last exercise, reliability was moderate-good for musculoskeletal radiologists and rheumatologists with previously demonstrated good scoring proficiency. Median pairwise single-measure/average-measure ICC were 0.67/0.80 for status scores and 0.69/0.82 for change scores; κ ranged 0.35–0.77.

Conclusion. Whole-body MRI scoring of peripheral arthritis and enthesitis is reliable, which encourages further testing and refinement in clinical trials. (J Rheumatol First Release April 15 2019; doi:10.3899/jrheum.181084)

Key Indexing Terms:
MAGNETIC RESONANCE IMAGING   OUTCOME ASSESSMENT   ARTHRITIS OMERACT

Magnetic resonance imaging (MRI) allows objective assessment of inflammation in peripheral joints and entheses[1,2,3,4,5,6,7]. MRI scoring systems have until now focused on assessing parts of the musculoskeletal system in detail, e.g., the Rheumatoid Arthritis MRI Scoring System (RAMRIS), which is applied to the wrist and metacarpophalangeal joints and adjacent tendon sheaths[8,9,10]. The interest in a whole-body MRI (WBMRI) approach is growing because modern MRI scanners permit whole-body scanning within an acceptable time frame (< 1 h), and future improvements in MRI hardware and pulse sequences are expected to improve scan time and image resolution further.

WBMRI of patients with inflammatory arthritis has mainly been investigated in small cross-sectional[2,5,6,7,11] or longitudinal studies[3,4,12]. To our knowledge, 2 randomized, controlled trials have used WBMRI as an outcome measure, applying different assessment systems[3,4]. To increase homogeneity, validity, and across-study comparability of WBMRI as outcome measure, the Outcome Measures in Rheumatology (OMERACT) MRI in the Arthritis Working Group decided to develop a scoring system for inflammation of peripheral joints and entheses for use in future phase II/III studies, which aim to objectively document the effect of an intervention on the inflammatory load in peripheral joints and entheses.

The objective was to develop an MRI Whole-Body Score for Inflammation in Peripheral Joints and Entheses in Inflammatory Arthritis (MRI-WIPE) and to investigate its feasibility and reliability.

MATERIALS AND METHODS

Development of the scoring system through iterative multireader scoring exercises. In 2016, the OMERACT MRI in Arthritis Working Group decided on inflammation in peripheral joints and entheses as the primary focus for WBMRI development, and then agreed on consensus MRI definitions for arthritis and enthesitis, selected anatomical locations for assessment, and a core set of MRI sequences and imaging planes for the different regions, and proposed a preliminary scoring system. It was decided to test and further develop the system by iterative multireader exercises[13,14,15].

In 2017–2018, 4 (3 cross-sectional and 1 longitudinal) Web-based multireader exercises were performed, separated by online training and calibration meetings. Schematics for recording the presence of lesions and their severity were drawn (SK and MØ; Figure 1). Subsequently, courtesy of CaRE

of Rheumatic and Musculoskeletal Medicine, University of Leeds, and NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospital NHS Trust; C.E. Althoff, MD, Department of Radiology, Arthritis Imaging Research Group, University Hospital Charité; J. Paschke, BSc, CaRE Arthritis; C. Peterfy, MD, PhD, Spire Sciences Inc.; K.A. Hermann, MD, PhD, Department of Radiology, Arthritis Imaging Research Group, University Hospital Charité; M. Østergaard, MD, PhD, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, and Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen.

Address correspondence to Dr. S. Krabbe, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Valdemar Hansens Vej 17, DK-2600 Glostrup, Denmark. E-mail: simonkrabbe@gmail.com

Accepted for publication January 24, 2019.

Arthritis, a Web-based schematic data entry interface was created (JP) and used together with a DICOM (Digital Imaging and Communication in Medicine) image viewer (Figure 2) to conduct entirely Web-based scoring exercises. In Exercise 1, 9 readers (1 radiologist, 8 rheumatologists) tested a draft scoring system in 2 patients with axial spondyloarthritis (axSpA). Results were discussed, and the system was slightly modified. In Exercise 2, 14 readers (3 radiologists, 11 rheumatologists) assessed 5 patients with axSpA. Discrepant cases and potential difficulties in applying the scoring system were discussed online to obtain consensus, train inexperienced readers, and identify potential pitfalls.

In Exercise 3, MRI of 8 patients [4 rheumatoid arthritis (RA), 4 psoriatic arthritis (PsA)] were scored by 14 readers (4 radiologists, 10 rheumatologists). Because of widely variable agreement (minimal-good) between reader pairs, 2 online meetings were held to improve calibration before proceeding to Exercise 4, in which MRI at 2 timepoints of 6 patients with axSpA who started tumor necrosis factor (TNF) inhibitor treatment were assessed by 10 readers (3 radiologists, 7 rheumatologists) blinded for chronology. In all exercises, readers were aware of the patient groups involved (SpA or RA), but not the diagnosis of individual cases.

Reader instructions containing definitions and image examples of normal findings (e.g., blood vessels) that could be mistaken for inflammation, and many examples of lesions with different grading were made available at www.copecare.dk and www.carearthritis.com. While Exercises 1 and 2 were used solely for qualitative training and understanding principles and pitfalls, for Exercises 3 and 4, reliability statistics were calculated (pairwise single measures and average measures intraclass correlation coefficients (ICC) by absolute agreement for sum scores and squared weights Cohen's κ for individual scores).

Approval was obtained from the Regional Committee on Health Research Ethics, Region Hovedstaden, Denmark (H-1-2013-118), and patients provided written informed consent.

Scoring methodology. Inflammation in joints (arthritis) and at entheses (enthesitis) are both assessed separately for soft tissues (synovitis at joints, soft tissue inflammation at entheses) and bone (osteitis), see Østergaard, *et al*[13] for exact MRI definitions.

Preferably, synovitis and soft tissue inflammation are assessed on T1-post-Gd images and osteitis on short-tau inversion recovery (STIR)/T2-weighted fat-sat (T2FS) images. But if STIR/T2FS is the only method available, synovitis and soft tissue inflammation can be assessed based on it. Each component is scored on a semiquantitative scale of 0–3 (none/mild/moderate/severe), following the principles from the RAMRIS and PsAMRIS systems[8,16]. In total, 83 peripheral joints and 33 entheses are assessed. The MRI-WIPE score is derived by adding all scores together; the total range is 0–738 (joints 0–537; entheses 0–201; Figure 2 and Appendix 1).

RESULTS

Readers from 10 different countries across the globe participated. Exercises 1 and 2 were used only for initial learning, calibration, and identification of pitfalls. In Exercise 3, agreement between readers varied from poor to good for the 4 lesion types and their sum scores (Table 1). Reliability varied between reader pairs depending on reader experience. When limiting the analysis to the 4 musculoskeletal radiologists, reliability improved to moderate-good.

The same pattern was observed in Exercise 4, where reliability was poor-good among all readers, but when restricted to the 3 musculoskeletal radiologists and 3 rheumatologists with the better reliability in the previous exercise, reliability was moderate-good. Thus, among the more trained readers, grading seemed reliable. MRI-WIPE reading time for 1 MRI was not measured but estimated to be ≤ 60 min.
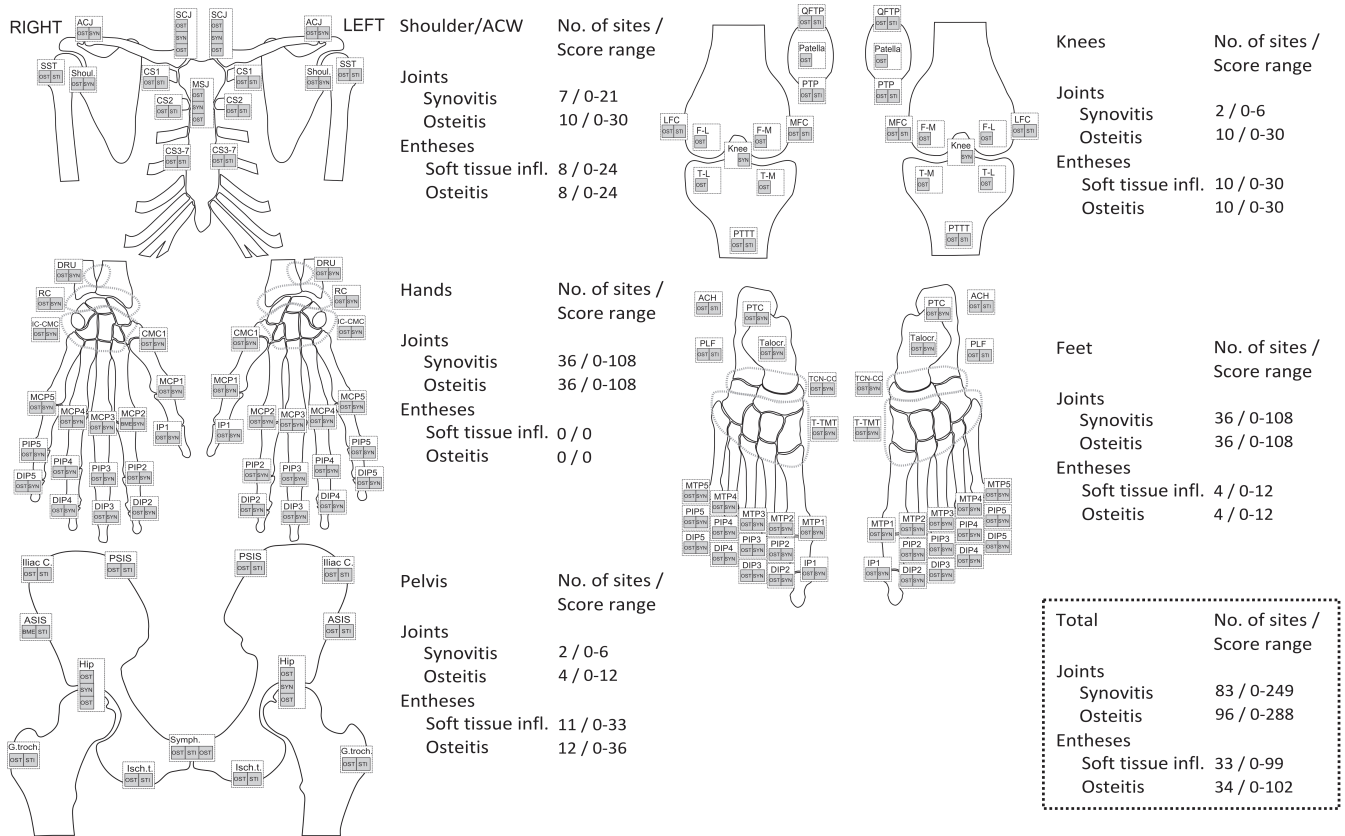
Figure 1. Data entry schematics and scoring ranges. Osteitis of the sternoclavicular joint is assessed separately for sternum and clavicle. Osteitis of the manubriosternal joint is assessed separately for manubrium and body of sternum. Osteitis of the hip joint is assessed separately for acetabulum and femur. Osteitis of the knee joint is assessed separately for lateral femur, medial femur, lateral tibia, medial tibia, and patella. Osteitis of the pubic symphysis is assessed separately for left and right pubic bone. OST: osteitis; SYN: synovitis; STI: soft tissue inflammation. Shoulder/ACW (anterior chest wall): ACJ: acromioclavicular joint; SCJ: sternoclavicular joint; SST: supraspinatus tendon; CS: costosternal joint; MSJ: manubriosternal joint; Should: glenohumeral joint. Hands: DRU: distal radioulnar joint; RC: radiocarpal joint; IC-CMC: intercarpal and carpometacarpal joints; MCP: metacarpophalangeal joint; PIP: proximal interphalangeal; DIP: distal interphalangeal. Pelvis: PSIS: posterior superior iliac spine; Iliac C: iliac crest; ASIS: anterior superior iliac spine; G troch: greater trochanter; Isch t: ischial tuberosity; Symph: pubic symphysis. Knees: QFTP: quadriceps femoris tendon insertion into patella; PTP: patellar tendon insertion into patella; PTTT: patellar tendon insertion into tibial tuberosity; MFC: medial femoral condyle; LFC: lateral femoral condyle; F-L: femur-lateral; F-M: femur-medial; T-L: tibia-lateral; T-M: tibia-medial. Feet: ACH: Achilles tendon; PLF: plantar fascia; PTC: posterior talocalcaneal joint; Talocr: talocrural joint; TCN-CC: talocalcaneon-avicular and calcaneocuboid joints; T-TMT: tarsal and tarsometatarsal joints; MTP: metatarsophalangeal.

Responsiveness of the MRI-WIPE score was good during TNF inhibitor treatment (mean change score –6.3, SD 6.5, and standardized response mean 1.0). Average-measure ICC based on 2 readers (status 0.80, change 0.82) were higher than single-measure ICC (status 0.67, change 0.69; Table 1). Using 3 readers, average-measure ICC were higher (status 0.86, change 0.86).

DISCUSSION

Definitions of key MRI pathologies and a scoring system (MRI-WIPE, MRI Whole-Body Score for Inflammation in Peripheral Joints and Entheses in Inflammatory Arthritis) were agreed upon by consensus in the OMERACT MRI in Arthritis Working Group. The scoring system was developed in analogy with RAMRIS/PsAMRIS scoring systems but allows assessment of multiple peripheral joints and entheses and is not limited to a specific diagnosis in its current form. In small cross-sectional and longitudinal reading exercises, the system had moderate-good reliability for status scores and change scores, when limiting the analysis to readers who were musculoskeletal radiologists or who had shown good proficiency of scoring in agreement with most readers in the previous exercises. Potentially, WBMRI could provide a high between-group discrimination in randomized controlled trials[4]. Thus, the scoring system appears promising for further validation and future use in randomized controlled trials.

A scan time of about 45 min for peripheral joints and entheses, and about 1 h if axial joints were included, was acceptable to the included patients. Thus, the approach was feasible, although no formal survey of patient satisfaction or discomfort was undertaken.

Subsequent steps may include tailoring/analyzing different
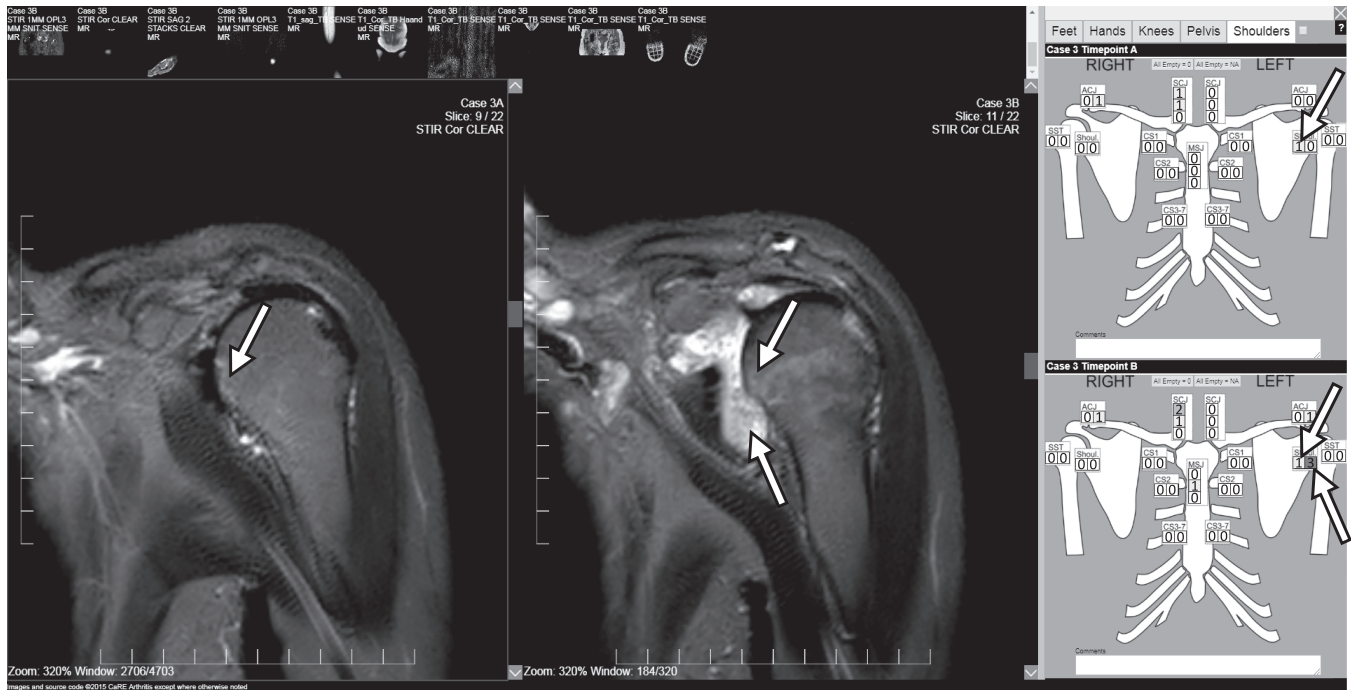
Figure 2. Web-based DICOM (Digital Imaging and Communication in Medicine) image viewer (provided courtesy of CaRE Arthritis at www.carearthritis.com). Short-tau inversion recovery images of the left shoulder region from the same patient at 2 timepoints (left side and middle) and the corresponding completed data entry schematics (right). White arrows: synovitis (score 3, severe) and osteitis (score 1, mild) of the left glenohumeral joint as assessed on the magnetic resonance images and entered in the corresponding data entry schematic.

joint combinations for different diseases (i.e., a modular approach, where only a selection of areas is imaged and scored, guided by the key questions in individual studies), because diseases such as RA, PsA, and axSpA have different patterns of joint and enthesis involvement. Analyzing different weighting of components, as recently attempted with the RAMRIS system[17], e.g., by putting less weight on small joints, may also be considered. Currently, WBMRI image quality is lowest in small joints because of their size and limited image resolution (slice thickness 3–5 mm), but new MRI units and sequence types can provide better resolution.

Not all readers reached the same level of reliability, but several readers' experiences in reading certain areas were also minimal, and as expected this could not be resolved by a few training exercises. Because of the complex anatomy and many regions to score, it is essential to use appropriate equipment, i.e., 1–2 large high-resolution monitors, in an appropriately lit room, where images of the needed number of timepoints are visualized in an appropriate size without zooming. An online training and calibration module, potentially with a final test of the reader's proficiency compared to expert readers, is a possibility. Investigating alternative MRI sequences or scanning protocols may also be an option.

Rather few cases were included in the exercises, but for the purposes of development, it was considered more important to understand and discuss potential discrepancies and try to calibrate readers. Higher patient numbers would have increased the certainty of the calculated reliability measures.

The MRI-WIPE score appears to be particularly reliable if the average score of 2 or 3 readers is used in the final analysis of a study, compared to scores based on only 1 reader, because the average measure ICC for 2 or 3 readers were substantially higher than single-measure ICC. With 3 readers, average measure ICC for status scores and for change scores were both 0.86.

The MRI-WIPE score is promising, because scoring was reliable between readers with previous good scoring proficiency. The system needs further validation in larger, longitudinal studies, but in its current form it could be of interest in trials striving for global measures of inflammation in peripheral joints and entheses.

REFERENCES
1. Eshed I, Bollow M, McGonagle DG, Tan AL, Althoff CE, Asbach P, et al. MRI of enthesitis of the appendicular skeleton in spondyloarthritis. Ann Rheum Dis 2007;66:1553-9.
2. Karpitschka M, Godau-Kellner P, Kellner H, Horng A, Theisen D, Glaser C, et al. Assessment of therapeutic response in ankylosing

Table 1. Interreader reliability.

Variables	No. Patients	Type of Score	MRI-WIPE (total score)	MRI-WIPE (subscore: osteitis, joints)	MRI-WIPE (subscore: synovitis, joints)	MRI-WIPE (subscore: osteitis, entheses)	MRI-WIPE (subscore: soft tissue inflammation, entheses)	ICC for Sum Scores (single measure)	ICC for Sum Scores (average measure)	κ at Lesion-level (osteitis, joints)	κ at Lesion-level (synovitis, joints)	κ at Lesion-level (osteitis, entheses)	κ at Lesion-level (soft tissue inflammation, entheses)
Exercise 3													
All 14 readers	8	Status scores	18.4 (10–32)	4.5 (2–9)	10.2 (5–21)	1.2 (0–3)	2.5 (1–5)	0.28 (−0.37 to 0.92)	0.44 (−1.18 to 0.96)	0.34 (0.05–0.74)	0.45 (0.21–0.71)	0.39 (0.00–1.00)	0.30 (−0.03 to 0.72)
4 musculoskeletal radiologists†	8	Status scores	16.8 (5–39)	4.3 (0–10)	9.3 (3–28)	0.8 (0–2)	2.5 (1–5)	0.67 (0.30–0.87)	0.80 (0.46–0.93)	0.51 (0.32–0.59)	0.59 (0.52–0.71)	0.47 (0.00–1.00)	0.40 (0.24–0.61)
Exercise 4													
All 10 readers	6	Status scores	15.1 (8–28)	4.0 (1–8)	5.2 (2–13)	3.1 (1–7)	2.7 (0–4)	0.55 (−0.08 to 0.96)	0.71 (−0.17 to 0.98)	0.49 (−0.02 to 0.83)	0.54 (0.17–0.84)	0.54 (−0.02 to 0.89)	0.28 (−0.03 to 0.72)
Same as above	6	Change scores	−5.2 (−10 to 3)	−1.3 (−3 to 1)	−1.2 (−5 to 2)	−1.8 (−5 to 0)	−0.8 (−3 to 1)	0.43 (−0.67 to 0.94)	0.60 (−4.0 to 0.97)	0.26 (−0.19 to 0.60)	0.22 (−0.12 to 0.69)	0.23 (−0.04 to 0.80)	0.14 (−0.09 to 0.66)
3 musculoskeletal radiologists‡ + 3 rheumatologists‡													
Status scores	6	Status scores	13.2 (5 to 26)	3.3 (0–9)	5.1 (1–13)	2.4 (0–7)	2.3 (0–5)	0.67 (0.37–0.96)	0.80 (0.54–0.98)	0.64 (0.29–0.83)	0.67 (0.53–0.84)	0.77 (0.57–0.89)	0.44 (0.19–0.72)
Same as above	6	Change scores	−6.3 (−14 to 2)	−1.8 (−4 to 1)	−1.9 (−7 to 1)	−1.7 (−5 to 0)	−1.0 (−4 to 0)	0.69 (0.27–0.94)	0.82 (0.43–0.97)	0.40 (−0.17 to 0.60)	0.40 (0.11–0.69)	0.42 (0.15–0.80)	0.35 (0.11–0.66)

Sum scores are mean (range) of the patients' scores (each patient's score is the average of the scores assigned to that patient). All ICC and κ values are median (range) of estimates for all reader pairs (1 reader pair for 2 readers, 3 reader pairs for 3 readers, etc.); ICC is 2-way model, single measure/average measure (with κ = 2 readers), by agreement; Cohen's κ with squared weights. Applied MRI sequences: T1-weighted precontrast images (exercise 1–4), post-contrast T1-weighted images (exercise 3), and STIR (exercise 1–4). MRI planes: shoulder/anterior chest wall: coronal; hands: coronal; pelvis: coronal; knees: sagittal; feet: sagittal of ankles and axial of feet (provides coronal view of feet). Readers: AJM, FG, IE†, MØ‡, PB, SJP‡, and RGL‡ (all exercises); DG and WPM (exercises 1–3); MSS and RGL‡ (exercises 2–4); RP (exercises 2–3); LJ† (exercises 3–4); JLJ† and VF (exercise 2); NH† (exercise 3). †Three rheumatologists with the better interreader reliability in previous scoring round. MRI-WIPE: Magnetic Resonance Imaging Whole-Body Score for Inflammation in Peripheral Joints and Entheses in Inflammatory Arthritis; STIR: short-tau inversion recovery; ICC: intraclass correlation coefficients.

spondylitis patients undergoing anti-tumour necrosis factor therapy by whole-body magnetic resonance imaging. Eur Radiol 2013;23:1773-84.

3. Song IH, Hermann KG, Haibel H, Althoff CE, Listing J, Burmester GR, et al. Effects of etanercept versus sulfasalazine in early axial spondyloarthritis on active inflammatory lesions as detected by whole-body MRI (ESTHER): A 48-week randomised controlled trial. Ann Rheum Dis 2011;70:590-6.

4. Krabbe S, Østergaard M, Eshed I, Sørensen IJ, Jensen B, Møller JM, et al. Whole-body magnetic resonance imaging in axial spondyloarthritis: reduction of sacroiliac, spinal, and entheseal inflammation in a placebo-controlled trial of adalimumab. J Rheumatol 2018;45:621-9.

5. Poggenborg RP, Pedersen SJ, Eshed I, Sørensen IJ, Møller JM, Madsen OR, et al. Head-to-toe whole-body MRI in psoriatic arthritis, axial spondyloarthritis and healthy subjects: first steps towards global inflammation and damage scores of peripheral and axial joints. Rheumatology 2015;54:1039-49.

6. Poggenborg RP, Eshed I, Østergaard M, Sørensen IJ, Møller JM, Madsen OR, et al. Enthesitis in patients with psoriatic arthritis, axial spondyloarthritis and healthy subjects assessed by 'head-to-toe' whole-body MRI and clinical examination. Ann Rheum Dis 2015;74:823-9.

7. Weckbach S, Schewe S, Michaely HJ, Steffinger D, Reiser MF, Glaser C. Whole-body MR imaging in psoriatic arthritis: additional value for therapeutic decision making. Eur J Radiol 2011; 77:149-55.

8. Østergaard 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.

9. Glinatsi D, Bird P, Gandjbakhch F, Haavardsholm EA, Peterfy CG, Vital EM, et al. Development and validation of the OMERACT rheumatoid arthritis magnetic resonance tenosynovitis scoring system in a multireader exercise. J Rheumatol 2017;44:1688-93.

10. Døhn UM, Conaghan PG, Eshed I, Boonen A, Boyesen P, Peterfy CG, et al. The OMERACT-RAMRIS rheumatoid arthritis magnetic resonance imaging joint space narrowing score: Intrareader and interreader reliability and agreement with computed tomography and conventional radiography. J Rheumatol 2014;41:392-7.

11. Axelsen MB, Eshed I, Duer-Jensen A, Møller JM, Pedersen SJ, Østergaard M. Whole-body MRI assessment of disease activity and structural damage in rheumatoid arthritis: first step towards an MRI joint count. Rheumatology 2014;53:845-53.

12. Axelsen MB, Eshed I, Østergaard M, Hetland ML, Møller JM, Jensen DV, et al. Monitoring total-body inflammation and damage in joints and entheses: the first follow-up study of whole-body magnetic resonance imaging in rheumatoid arthritis. Scand J Rheumatol 2017;46:253-62.

13. Østergaard M, Eshed I, Althoff CE, Poggenborg RP, Diekhoff T, Krabbe S, et al. Whole-body magnetic resonance imaging in inflammatory arthritis: systematic literature review and first steps toward standardization and an OMERACT scoring system. J Rheumatol 2017;44:1699-705.

14. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745-53.

15. Boers M, Kirwan JR, Tugwell P, Beaton D, Bingham CO III, Conaghan PG, et al. The OMERACT Handbook. [Internet. Accessed September 6, 2018.] Available from: https://omeract.org/resources

16. Østergaard M, McQueen F, Wiell C, Bird P, Bøyesen P, Ejbjerg B, et al. The OMERACT Psoriatic Arthritis Magnetic Resonance Imaging Scoring System (PsAMRIS): Definitions of key pathologies, suggested MRI sequences, and preliminary scoring system for PsA hands. J Rheumatol 2009;36:1816-24.

17. Sundin U, Østergaard M, Glinatsi D, Aga AB, Hørslev-Petersen K, Hetland ML, et al. Validity and responsiveness of combined inflammation and combined joint damage scores based on the OMERACT Rheumatoid Arthritis MRI Scoring System (RAMRIS). J Rheumatol 2019 Feb 15 (E-pub ahead of print).

Further Details on the Scoring Methodology	
Osteitis	Osteitis should be assessed in the bone from the articular surface/entheseal insertion to a depth of 1 cm on all available images. Grading scale: The scale is 0-3 based on the proportion of bone with edema, compared to the "assessed bone volume", judged on all available images: 0: no edema; 1: 1-33% of bone edematous; 2: 34-66% of bone edematous; 3: 67-100%.
Synovitis	Synovitis should be assessed in the entire synovial compartment on all available images. Grading scale: Score 0 is normal, while 1-3 is mild, moderate, severe, by thirds of the maximum potential volume of enhancing tissue in the synovial compartment.
Soft tissue inflammation	Soft tissue inflammation should be assessed inside the dense fibrous connective tissue part of the enthesis (which is continuous with and indistinguishable from the ligament/tendon) and in its immediate surroundings to a distance of 1 cm from the entheseal insertion. Grading scale: Score 0 is normal, while 1-3 is mild, moderate, severe, by thirds of the maximum potential volume of enhancing tissue. A distance of 1 cm was chosen by consensus in the OMERACT MRI in Arthritis Working Group with the aim of capturing inflammation that originates from the enthesis and not capturing tendinopathy.
Positive vs. negative score	A positive score of 1 should only be made when the reader is confident that there is an abnormality. All synovial joints contain normal joint fluid; this should not be scored. The scoring system aims at scoring inflammation. If the reader is hesitating whether to score a possible lesion 1 (mild) or 0 (none), it should probably be scored 0 (none).
Lesion judged borderline between two scores	If the lesion is judged borderline 1 vs. 2 or 2 vs. 3, lesion intensity may be taken into account. E.g. if a lesion is borderline between 1 (mild) and 2 (moderate), it may be scored 1 (mild) if not judged intense. Similarly, e.g. if a lesion is borderline between 2 (moderate) and 3 (severe), it may be scored 3 (severe) if judged intense. When there is an increased amount of synovial tissue, not just effusion, and the lesion is judged borderline between two scores, the higher score may be assigned.
No detailed rules for scoring each specific joint or enthesis	To allow a feasible scoring, we did not introduce detailed rules for how to score each specific joint or enthesis, as we aimed at only having generic rules, e.g. to assess soft tissue changes until 1 cm from insertional site irrespective of enthesis. Therefore, if e.g. the retrocalcaneal bursa is inflamed and it partly lies within 1 cm of the Achilles tendon insertion, that part should be considered when assessing soft tissue inflammation.
Choice of MRI sequences	Preferentially, synovitis and soft tissue inflammation are assessed on T1-post-Gd images and osteitis on Short Tau Inversion Recovery (STIR)/T2-Weighted Fat-Sat (T2FS) images, but if only STIR/T2FS is available, synovitis and soft tissue inflammation can be assessed based on this.

List of Sites Assessed		
JOINTS	No. Sites (synovitis)	No. Sites (osteitis)
Acromioclavicular joint	2	2
Sternoclavicular joint†	2	4
Manubriosternal joint†	1	2
Glenohumeral joint	2	2
Distal radioulnar joint	2	2
Radiocarpal joint	2	2
Intercarpal/carpometacarpal joints 2-5	2	2
Carpometacarpal joint 1	2	2
Metacarpophalangeal joints 1-5	10	10
Interphalangeal joint 1 (hands)	2	2
Proximal interphalangeal joints 2-5 (hands)	8	8
Distal interphalangeal joints 2-5 (hands)	8	8
Hip joint†	2	4
Knee joint†	2	10
Talocrural joint	2	2
Posterior talocalcaneal joint	2	2
Talocalcaneonavicular/calcaneocuboid joints	2	2
Tarsal/tarsometatarsal joints	2	2
Metatarsophalangeal joints 1-5	10	10
Interphalangeal joint 1 (feet)	2	2
Proximal interphalangeal joints 2-5 (feet)	8	8
Distal interphalangeal joints 2-5 (feet)	8	8
TOTAL NO. OF SITES	83	96
SCORE RANGE	0-249	0-288
ENTHESES	No. Sites (soft tissue infl)	No. Sites (osteitis)
Supraspinatus tendon	2	2
Costosternal joint 1	2	2
Costosternal joint 2	2	2
Costosternal joint 3-7	2	2
Posterior superior iliac spine	2	2
Iliac crest	2	2
Anterior superior iliac spine	2	2
Ischial tuberosity	2	2
Pubic symphysis†	1	2
Greater trochanter	2	2
Quadriceps femoris tendon insertion into patella	2	2
Patellar tendon insertion into patella	2	2
Patellar tendon insertion into tibial tuberosity	2	2
Medial femoral condyle	2	2
Lateral femoral condyle	2	2
Achilles tendon	2	2
Plantar fascia	2	2
TOTAL NO. SITES	33	34
SCORE RANGE	0-99	0-102

†Osteitis of the sternoclavicular joint is assessed separately for sternum and clavicle. Osteitis of the manubrosternal joint is assessed separately for manubrium and body of sternum. Osteitis of the hip joint is assessed separately for acetabulum and femur. Osteitis of the knee join is assessed separately for lateral femur, medial femur, lateral tibia, medial tibia, and patella. Osteitis of the pubic symphysis is assessed separately for left and right pubic bone.