Title

The OMERACT MRI in Enthesitis Initiative: Definitions of Key Pathologies, Suggested MRI Sequences and Novel Heel Enthesitis Scoring System (HEMRIS)

Author Name	ORCID ID
Ashish J Mathew	https://orcid.org/0000-0002-2061-2042
Simon Krabbe	https://orcid.org/0000-0002-2877-1582
Iris Eshed	https://orcid.org/0000-0002-4655-9606
Frédérique Gandjbakhch	https://orcid.org/0000-0002-3181-4598
Paul Bird	https://orcid.org/0000-0003-3314-3270
Susanne J Pedersen	https://orcid.org/0000-0002-6500-9263
Maria S Stoenoiu	https://orcid.org/0000-0003-2107-8846
Violaine Foltz	https://orcid.org/0000-0003-2097-9239
Daniel Glinatsi	https://orcid.org/0000-0003-0452-8200
Robert G Lambert	https://orcid.org/0000-0003-0305-3290
Kay Geert A Hermann	https://orcid.org/0000-0001-6142-3814
Walter P Maksymowych	https://orcid.org/0000-0002-1291-1755
Ida K Haugen	https://orcid.org/0000-0001-7810-2216
Jacob L Jaremko	https://orcid.org/0000-0001-5314-2297
René P Poggenborg	
Joel Paschke	
Jean-Denis Laredo	https://orcid.org/0000-0002-5813-0268
Philippe Carron	https://orcid.org/0000-0001-9254-6171
Philip G. Conaghan	https://orcid.org/0000-0002-3478-5665
Mikkel Østergaard	https://orcid.org/0000-0003-3690-467X

Key indexing terms

Enthesopathy, Magnetic resonance imaging, Spondyloarthropathy, Psoriatic arthritis, OMERACT

Departments/Institutions

Department of Clinical Immunology & Rheumatology, Christian Medical College, Vellore, India; Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark; Department of Diagnostic Imaging, Sheba Medical Center, Affiliated to the Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel; Hôpitaux Universitaires Pitié Salpêtrière, Paris, France; Paris 6 University, GRC-UPMC 08, Pierre Louis Institute of Epidemiology and Public Health, Paris, France; Division of Medicine, University of New South Wales, Sydney, Australia; Cliniques Universitaires Saint-Luc, Institut de Recherche Expérimentale et Clinique (IREC), Université catholique de Louvain, Brussels, Belgium; Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, Canada; Department of Radiology, Arthritis Imaging Research Group, University Hospital Charité, Berlin, Germany; CaRE (Canadian Research Education) Arthritis and Department of Medicine, University of Alberta, Edmonton, Canada; Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway; CaRE (Canadian Research Education) Arthritis, Edmonton, Canada; Service de Radiologie, Hôpital Lariboisiére, APHP & Université Paris-Diderot, Paris, France; Department of Rheumatology, Ghent University Hospital, Ghent, Belgium; Faculty of Medicine and Health, University of Leeds, United Kingdom

Source of financial support

This project has not received any outside funding

Conflicts of interest

None

Authors

AJ. Mathew, MB BS, DNB, DM, Associate Professor, Department of Clinical Immunology and Rheumatology, Christian Medical College, Vellore, India; PhD Fellow, Department of Clinical Medicine, Faculty of Health and Medical Sciences, COPECARE, Center for Rheumatology and Spine Diseases, Rigshospitalet Glostrup, University of Copenhagen, Copenhagen, Denmark

S. Krabbe, MD, PhD Fellow, Department of Clinical Medicine, Faculty of Health and Medical Sciences, COPECARE, Center for Rheumatology and Spine Diseases, Rigshospitalet Glostrup, University of Copenhagen, Copenhagen, Denmark
I. Eshed, MD, Professor, Department of Diagnostic Imaging, Sheba Medical Center, Affiliated to the Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel
F. Gandjbakhch, MD, Practising Rheumatologist, Hôpitaux Universitaires Pitié Salpêtrière, Paris 6 University, GRC-UPMC 08, Pierre Louis Institute of Epidemiology and Public Health, Paris, France
P. Bird, B Med (Hons), FRACP, PhD, Grad Dip MRI, Associate Professor, University of NSW, Sydney, Australia
SJ. Pedersen, MD, PhD, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark
MS. Stoenoiu, MD, PhD, Cliniques Universitaires Saint-Luc, Institut de Recherche Expérimentale et Clinique (IREC), Université catholique de Louvain, Brussels, Belgium
V. Foltz, MD, Practising Rheumatologist, Hôpitaux Universitaires Pitié Salpêtrière, Paris, France
D. Glinatsi, MD, PhD, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark
RG. Lambert, MB.BCh, FRCR, FRCPC, Professor, Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, Canada
KGA. Hermann, MD, PhD, Senior Consultant, Department of Radiology, Arthritis Imaging Research Group, University Hospital Charité, Berlin, Germany
WP. Maksymowych MD, FRCP(C), Professor, Division of Rheumatology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Canada
IK. Haugen, MD, PhD, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway
JL. Jaremko, MD, PhD, FRCPC, Associate Professor, Department of Radiology, University of Alberta, Edmonton, Canada
RP. Poggenborg, MD, PhD, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Denmark
J, Paschke, BSc, CaRE (Canadian Research Education) Arthritis, Alberta, Edmonton, Canada
JD. Laredo, MD, Professor, Service de Radiologie, Hôpital Lariboisiére, APHP & Université Paris-Diderot, Paris, France
P. Carron, MD, PhD, Department of Rheumatology, Ghent University Hospital, Ghent, Belgium
PG. Conaghan MB BS, PhD, FRACP, FRCP, Professor of Musculoskeletal Medicine, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, & NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, United Kingdom
M. Østergaard, MD, PhD, DMSc, Professor, Department of Clinical Medicine, Faculty of Health and Medical Sciences, COPECARE, Center for Rheumatology and Spine Diseases, Rigshospitalet Glostrup, University of Copenhagen, Copenhagen, Denmark

Correspondence

Mikkel Østergaard, MD, PhD, DMSc, COPECARE – Center for Rheumatology and Spine Diseases, Rigshospitalet, Nordre Ringvej 57, DK-2600 Glostrup, Denmark
Email: mo@dadlnet.dk

Running head

OMERACT heel enthesitis MRI score

Word count
1500/1500

Abstract

Objectives: To develop and validate an enthesitis MRI-scoring system for spondyloarthritis/psoriatic arthritis, using the heel as model.

Methods: Consensus definitions of key pathologies and three heel enthesitis multi-reader scoring exercises were done, separated by discussion, training and calibration.

Results: Definitions for bone and soft tissue pathologies were agreed. In final exercise, median pairwise single-measures intra-class correlation coefficients(ICCs; patient-level) for entheseal inflammation status/change scores were 0.83/0.82 for all readers. For radiologists and selected rheumatologists ICCs were 0.91/0.84 and quadratic-weighted kappas(lesion-level) 0.57-0.91/0.45-0.81.

Conclusion: The proposed definitions and heel enthesitis scoring system (HEMRIS) are reliable among trained readers and promising for clinical trials.

Total words: 100

Introduction

Enthesitis, inflammation at insertion sites of ligaments, fasciae, tendons and joint capsules to bone, is a central feature of spondyloarthritis (SpA), including psoriatic arthritis (PsA). Sensitive and objective assessment of enthesitis is important in SpA clinical trials. Conventional clinical methods have limited reliability, validity and sensitivity[1-3]. Magnetic resonance imaging (MRI) is a sensitive method for detecting enthesitis in peripheral SpA and the only method allowing detection of peri-entheseal osteitis[4-6]. MRI studies have demonstrated decreased entheseal inflammation after anti-tumor necrosis factor (TNF) therapy, but no validated MRI-scoring systems exist for evaluating enthesitis in clinical trials [7]. Our aim was to create consensus-based MRI-definitions of key enthesitis pathologies and through multi-reader exercises to develop and validate an MRI score for assessing enthesitis in patients with SpA, focusing on the heel region.

Methods

The OMERACT MRI in Arthritis Working Group initially performed a systematic literature review (SLR) of studies with MRI being used for assessment of enthesitis [8]. Based on this SLR, MRI-sequences for optimal visualization of enthesitis were identified, and MRI-definitions of key enthesitis pathologies were decided by consensus between group members through meetings/e-mails. The heel region (insertions of Achilles tendon and plantar fascia) was chosen for initial testing due to its frequent involvement. Three multi-reader exercises, with consensus discussion and calibration in-between were then performed. A graphical data entry schematic (Appendix-Figure 1A) was created, and subsequently a web -based interface which simultaneously displayed DICOM-images and the data entry schematic (Appendix-Figure 1B). In Exercise 1, performed to identify challenges and pitfalls, sagittal T1-weighted (T1w) and sagittal and axial T2w-fat-suppressed (T2wFS) MR-images of 10 ankles (4 inflammatory enthesitis (peripheral SpA), 4 mechanical enthesitis and 2 normal controls) were scored by 15 readers from 10 countries), with varying expertise in ankle MRI, for enthesitis at Achilles tendon and plantar fascia insertions. This was followed by a web-based calibration exercise leading to minor score sheet modifications. In Exercise 2, 16 ankle MRIs (8 inflammatory enthesitis (peripheral SpA), 3 mechanical enthesitis and 5 normal controls; MRI-sequences as above) were scored by 16 readers. In Exercise 3, ankle MRIs (sagittal T2wFS only) of 21 SpA patients from a clinical trial, obtained before and after anti-TNF therapy, were scored for inflammatory pathologies by 10 readers, blinded to chronological order. For assessing the reliability scores among the more experienced readers, agreement between the participating radiologists and the 3 rheumatologists with best overall ICCs for inflammatory pathologies in exercise 2 were analyzed separately.

Statistical analysis

Exercise 1 was mainly used for qualitative training and understanding principles and pitfalls, while for Exercises 2-3 reliability statistics (pairwise single measures and average measures intraclass correlation coefficients by absolute agreement (smICC and amICC) for sum scores (patient level) and squared weights Cohen's kappa for individual component scores (lesion level) were calculated. In Exercise 3, the standardized response mean (SRM) was calculated.

Results

Definitions of key pathologies

Key entheseal pathologies were selected and their definitions agreed upon by consensus within the OMERACT MRI in inflammatory arthritis working group (Table 1), based on knowledge from an SLR[8], and published OMERACT MRI definitions for comparable conditions[9-11]. The selected pathologies were intra-tendon hypersignal (entheseal tendonitis), peri-tendon hypersignal (entheseal peritendinitis), bone marrow edema (entheseal osteitis), bursitis, tendon thickening, enthesophyte, entheseal bone erosion and intra-tendon hypersignal on T1w sequence.

MRI sequences and planes

For evaluating inflammatory pathologies, it was agreed to include a fluid-sensitive sequence (short-tau inversion recovery(STIR) or T2wFS), and/or a fat-suppressed T1w-sequence following intravenous gadolinium (Gd) injection *(see appendix)*. A T1w-sequence prior to contrast injection (T1-pre-Gd) was considered helpful in determining the exact localization of inflammatory pathologies, due to its high anatomical resolution, and is essential for assessment of structural pathologies.

Scoring system

It was decided to score all assessed pathologies on a semiquantitative scale of 0-3 (none/mild/moderate/severe), following the principles from the RAMRIS and PsAMRIS systems[9-11], and to create a total entheseal inflammation score by summation of scores of all inflammatory parameters (intra-tendon hypersignal on T2w/STIR sequences, peri-tendon hypersignal, bone marrow edema and bursitis). Similarly, a total entheseal structural damage score by summation of structural scores (enthesophyte, bone erosion, tendon thickening) was evolved. Intra-tendon hypersignal on T1w sequences was not included in sum scores. In exercises described in the present paper, scoring of entheses of the heel region was chosen, i.e. at calcaneal insertions of the Achilles tendon and plantar fascia, respectively.

Exercise 1

Exercises 1 and 2 included single-point images of the heel region, which were scored for the selected pre-defined pathologies. Exercise 1 was used for initial learning, calibration and identification of pitfalls. Mean pairwise inter-reader single-measure ICCs for inflammatory and structural variables, done without calibration, were 0.40 and 0.41, respectively.

Exercise 2

In Exercise 2, agreement between reader pairs varied from poor-very good for various lesion types and their sum scores (Table 2). When limiting the analyses to three participating musculoskeletal radiologists and three rheumatologists with best ICCs for inflammatory pathologies in exercise 2, reliability improved to moderate-very good. For this subset of readers, median single-measures ICCs for total inflammation scores were 0.85, while for total structural damage scores 0.68. Median kappas for different inflammatory pathologies varied from 0.60-0.89, and for individual structural pathologies from 0.41-0.78. Average-measure ICCs based on two readers among the pre-selected 6 readers (median 0.92 for total inflammatory score, 0.81 for total damage scores) were better than the above-mentioned single-measure ICCs.

Exercise 3

This exercise included two-time point images, in which inflammatory pathologies were scored. Mean pairwise inter-reader ICCs and lesion-wise kappa agreement demonstrated moderate to good reliability when all readers were considered (Table 3). The subset of readers (3 rheumatologists with best agreement for inflammatory parameters in exercise 2 and the participating radiologist in exercise 3) demonstrated good to very good reliability, both for baseline scores and for change in scores (Table 2); the median baseline single-measures ICCs for total inflammation was 0.91, while 0.84 for change in score. Median average-measure ICCs based on two readers (status: 0.95(range 0.95-0.97); change: 0.92(0.89-0.96)) were higher than single-measure ICCs. Using three readers demonstrated numerically higher average-measure ICCs (status: median 0.97(0.97-0.97); change 0.94(0.94-0.95)).

HEMRIS showed moderate responsiveness, with SRM of 0.70(95%CI 0.38-1.05) for all readers in exercise 3.

Discussion

This study is the first international consensus effort towards development of a comprehensive MRI-scoring system, combined with MRI definitions and reader rules, for enthesitis in patients with spondyloarthritides. The work was informed by a SLR[8], which clarified knowledge gaps and need for development of a validated MRI enthesitis scoring system to be used as outcome measure in clinical trials. Enthesitis, often located at heels is a typical feature of SpA and is easily accessible for MRI[12]. Furthermore, enthesitis in SpA may show changes both in inflammation (such as bone marrow edema and peri-entheseal inflammation) and damage (such as erosion and new bone formation)[13,14]. Thus, both inflammatory and structural MRI findings were considered relevant to include in the scoring system. A series of multireader scoring exercises focused on the heel region, using an intuitive web-based data entry and image display platform. The preliminary heel enthesitis scoring system (OMERACT-HEMRIS) showed good inter-reader agreement for status scores and for change over time in inflammatory parameters. Considering that baseline heel enthesitis was not mandatory in exercise 3, the moderate SRM (0.70) supports that responsiveness of the HEMRIS score would likely be good in trials with baseline enthesitis as an inclusion criterion. Thus, HEMRIS appears promising for further validation and future use in randomized controlled trials.

The strengths of this initiative include taking a SLR as starting point to clarify unmet need, the involvement of experienced MRI researchers in the development of consensus-based definitions and scoring systems, the participation of multiple readers with both radiological and rheumatological backgrounds in interactive web-based exercises with standardized image display and scoring module. Limitations include varying experience and backgrounds of readers in the exercises which needs to be taken into consideration when interpreting the results. This was addressed by sub-analysis of scores of a subset of experienced readers, who had showed high scoring proficiency in previous exercises. Longitudinal studies incorporating T1w images are needed for assessment of the sensitivity to change of structural parameters. Future developments should also include an MRI enthesitis reference image atlas, and image sets for training and calibration. The definitions and scoring principle may be applicable to other entheses. Thus, validation of the definitions and scoring system in other anatomical regions are also suggested.

The heel enthesitis MRI score appears to be particularly reliable if the mean score of two readers (compared to one) is used in the final study analysis; the average measure ICCs for 2 readers were markedly higher (0.92-0.95 for inflammation total status/change score in last exercise) than single measure ICCs. This will be relevant in real life clinical trials where two independent readers generally score images.

Increasing novel therapeutic options in SpA and PsA increases the potential utility of an objective and reproducible enthesitis outcome measure. The proposed OMERACT MRI heel enthesitis scoring system (HEMRIS) is a promising tool for further refinement and validation through the OMERACT filter and for future use in clinical trials[15,16].

Acknowledgements

We thank the Canadian Research and Education (CaRE) Arthritis (www.carearthritis.com) for help with organization of online meetings and exercises and development of the web-based scoring interface.

PGC is supported in part by the UK National Institute for Health Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References:

1. Schett G, Lories RJ, D'Agostino MA, Elewaut D, Kirkham B, Soriano ER, et al. Enthesitis: from pathophysiology to treatment. Nat Rev Rheumatol 2017;13:731-41
2. Polachek A, Li S, Chandran V, Gladman DD. Clinical enthesitis in a prospective longitudinal psoriatic arthritis cohort: incidence, prevalence, characteristics and outcome. Arthritis Care Res 2017;69:1685-91
3. Mease PJ. Measures of psoriatic arthritis: Tender and Swollen Joint Assessment, Psoriasis Area and Severity Index (PASI), Nail Psoriasis Severity Index (NAPSI), Modified Nail Psoriasis Severity Index (mNAPSI), Mander/Newcastle Enthesitis Index (MEI), Leeds Enthesitis Index (LEI), Spondyloarthritis Research Consortium of Canada (SPARCC), Maastricht Ankylosing Spondylitis Enthesis Score (MASES), Leeds

Dactylitis Index (LDI), Patient Global for Psoriatic Arthritis, Dermatology Life Quality Index (DLQI), Psoriatic Arthritis Quality of Life (PsAQOL), Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F), Psoriatic Arthritis Response Criteria (PsARC), Psoriatic Arthritis Joint Activity Index (PsAJAI), Disease Activity in Psoriatic Arthritis (DAPSA), and Composite Psoriatic Disease Activity Index (CPDAI). Arthritis Care Res 2011;63 Suppl 11:S64-85

4. 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;66:1553-50
5. Mathew AJ, Coates LC, Danda D, Conaghan PG. Psoriatic arthritis: lessons from imaging studies and implications for therapy. Exp Rev Clin Immunol 2017;13:133-42
6. Benjamin M, Moriggl B, Brenner E, Emery P, McGonagle D, Redman S. The 'enthesis organ' concept: why enthesopathies may not present as focal insertional disorders. Arthritis Rheum 2004;50:3306-13
7. Dougados M, Combe B, Braun J, Landwew R, Sibilia J, Cantagrel A, et al. A randomised, multicentre, double-blind, placebo-controlled trial of etanercept in adults with refractory heel enthesitis in spondyloarthritis: the HEEL trial. Ann Rheum Dis 2010;69:1430-35
8. Mathew AJ, Krabbe S, Kirubakaran R, Barr AJ, Conaghan PG, Bird P, Østergaard M. Utility of magnetic resonance imaging (MRI) in diagnosing and monitoring enthesitis in patients with spondyloarthritis: a systematic literature review. J Rheumatol: (submitted as OMERACT article, September 2018)
9. 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 scoring system (OMERACT 2002 RAMRIS).
   J Rheumatol 2003;30:1385-86
10. 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

11. Østergaard M, Peterfy CG, Bird P, Gandjbakhch F, Glinatsi D, Eshed I, et al. The OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging (MRI) Scoring system: Updated recommendations by the OMERACT MRI in Arthritis Working Group.
J Rheumatol 2017;44:1706-12
12. Dougados M, van der Linden S, Juhlin R, Huitfledt B, Amor B, Calin A, et al. The European Spondyloarthropathy Study Group preliminary criteria for the classification of spondyloarthropathy. Arthritis Rheum 1991;34:1218-27
13. Feydy A, Lavie-Brion MC, Gossec L, Lavie F, Guerini H, Nguyen C, et al. Comparative study of MRI and power Doppler ultrasonography of the heel in patients with spondyloarthritis with and without heel pain and in controls. Ann Rheum Dis 2012;71:498-503
14. Kehl AS, Corr M, Weisman MH. Enthesitis – New insights into pathogenesis, diagnostic modalities and treatment. Arthritis Rheumatol 2016;68:312-22
15. Boers M, Kirwan JR, Tugwell P, Beaton D, Bingham CO III, Conaghan PG, et al. The OMERACT Handbook. (Internet. Accessed September 1, 2018). Available from: www.omeract.org/pdf/OMERACT_Handbook.pdf
16. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measuresment sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745-53

Table 1: MRI definitions of key entheseal pathologies		
	Pathology	**Definition**
1	Intra-tendon hypersignal (entheseal tendonitis)	Signal characteristics consistent with increased water content/inflammation* within the tendon/ ligament/aponeurosis close to its insertion
2	Peri-tendon hypersignal (entheseal peritendinitis)	Signal characteristics consistent with increased water content/inflammation * in the soft tissues surrounding the tendon/ ligament/aponeurosis, close to its insertion
3	Bone marrow edema (entheseal osteitis)	Bone lesion with ill-defined margins and signal characteristics consistent with increased water content/inflammation*, close to the tendon/aponeurosis insertion
4	Bursitis†	Signal characteristics consistent with increased water content/inflammation* in an above-normal sized bursa
5	Tendon/aponeurosis thickening	Abnormal thickening of the tendon/aponeurosis close to its insertion
6	Enthesophyte	Abnormal bone formation at the insertion of tendon/ ligament/aponeurosis insertion into the bone
7	Bone erosion (entheseal bone erosion)	A sharply marginated bone lesion, with typical signal characteristics** and a visible cortical break, located close to the tendon/ ligament/aponeurosis insertion
8	Intra-tendon hypersignal on T1w	Increased signal in T1-weighted sequence within the tendon/ ligament/aponeurosis close to its insertion

† This lesion should only be assessed in entheseal regions in which a relevantly located bursa is present.

* High signal intensity on STIR/T2wFS images and/or above normal post-gadolinium enhancement on T1w images

** On T1w images without contrast injection: loss of normal low signal intensity of cortical bone and loss of normal high signal intensity of marrow fat.

Table 2: Exercise 2: Single measure Inter-reader ICCs (sum scores), quadratic weighted kappa (individual component scores, per lesion) and mean scores of all readers

INFLAMMATORY PATHOLOGIES

	Reproducibility (smICC & Kappa)					Reader scores	
	All readers		Subset of readers *		Range of scores	All readers	Subset of readers *
	Median	Mean (Range)	Median	Mean (Range)		***Mean (Range)	Mean (Range)
Inter-reader smICC (patient level)							
Total inflammation scores	0.58	0.56 (0.11-0.90)	0.85	0.83 (0.76–0.90)	**0-21**	3.94 (0.67–8.27)	4.83 (0.5–10.67)
Inter-reader quadratic weighted kappa (lesions level)							
Achilles tendon							
Peri-tendon hypersignal	0.41	0.45 (0.14–1.00)	0.66	0.64 (0.28–0.89)	**0-3**	0.45 (0 – 2.00)	0.64 (0 – 2.50)
Intra-tendon hypersignal	0.50	0.47 (0.04–0.91)	0.68	0.71 (0.53–0.90)	**0-3**	0.62 (0.07 – 2.07)	0.73 (0 – 2.17)
Retro-calcaneal bursitis	0.47	0.45 (-0.06-0.86)	0.60	0.62 (0.47–0.71)	**0-3**	0.43 (0 – 1.67)	0.50 (0 – 2)
Bone marrow edema	0.83	0.78 (0.26-1.00)	0.89	0.90 (0.83-1.00)	**0-3**	0.44 (0 – 2.27)	0.52 (0 – 2.50)
Plantar fascia							
Peri-aponeurosis hypersignal	0.67	0.63 (0.12–0.91)	0.83	0.83 (0.74–0.91)	**0-3**	0.82 (0 – 2.53)	1.02 (0 – 3.00)
Intra-aponeurosis hypersignal	0.45	0.40 (0 – 0.92)	0.70	0.69 (0.54–0.92)	**0-3**	0.51 (0 – 1.60)	0.69 (0 – 2.33)
Bone marrow edema	0.84	0.77	0.86	0.86 (0.73-0.94)	**0-3**	0.66 (0-2.47)	0.74 (0-2.67)

		(0.11-0.98)						
STRUCTURAL PATHOLOGIES								
Inter-reader smICC (patient level)								
Total structural damage score	0.27	0.35 (-0.04-0.85)	0.68	0.66 (0.37–0.85)		**0-18**	1.54 (0.2 – 4.4)	2.33 (0.33 – 7.00)
Inter-reader quadratic weighted kappa (lesion level)								
Achilles tendon								
Tendon thickness	0.52	0.48 (0 – 0.92)	0.76	0.72 (0.41–0.92)		**0-3**	0.54 (0 – 2.27)	0.78 (0 – 3.00)
Bone erosion	0.54	0.45 (0 – 1.00)	0.78	0.78 (0.52–1.00)		**0-3**	0.14 (0 – 1.4)	0.19 (0 – 1.83)
Bone spur	0.00	0.26 (-0.08–1.0)	0.41	0.37 (0 – 0.87)		**0-3**	0.13 (0 – 0.87)	0.22 (0 – 1.33)
Intra-tendon hypersignal on T1w[@]	0.30	0.33 (-0.09-0.88)	0.64	0.63 (0.36–0.96)		**0-3**	0.46 (0.07 – 1.47)	0.58 (0 – 2.00)
Plantar fascia								
Tendon thickness	0.31	0.35 (-0.23-0.97)	0.86	0.72 (0.26–0.97)		**0-3**	0.50 (0 – 1.53)	0.75 (0 – 2.5)
Bone erosion	0.00	0.02 (-0.17-0.64)	0.00	0.03 (-0.05-0.14)		**0-3**	0.06 (0 – 0.27)	0.11 (0 – 0.5)
Bone spur	0.00	0.12 (-0.18-0.76)	0.42	0.4 (-0.18-0.76)		**0-3**	0.17 (0 – 0.53)	0.28 (0 – 1.17)
Intra-aponeurosis hypersignal on T1w[@]	0.21	0.25 (-0.19-0.84)	0.40	0.42 (0.05–0.83)		**0-3**	0.13 (0 – 1.27)	0.49 (0 – 2.00)

smICCs: - single measures intraclass correlation coefficient by two-way random effects, absolute agreement for sum scores (patient level).

* Three participating radiologists + three rheumatologists with best individual ICCs with other readers for inflammatory pathologies in exercise 2. ***Each patient's score was calculated as the mean of all readers. The presented mean and ranges are means/ranges of these values. (Range of readers' mean scores)

Readers: AJM, DG, FG, IH, IE †, KGH †, MS, MØ ‡, PB, RGL †, SK ‡, SJP, VF ‡, WM, (Exercises 1 and 2); JJ † (Only Exercise 1); RPP (Only Exercise 2). † Musculoskeletal radiologists. ‡ Three rheumatologists with best individual ICC for inflammatory pathologies in exercise 2.

@ Not included in total structural damage score; it may occur both on inflammatory and structural backgrounds.

Table 3: Exercise 3: Baseline and change single-measure inter-reader ICCs (sum scores), quadratic weighted kappa (individual component scores, per lesion) and mean scores of all readers

	Reproducibility (smICC & Kappa)					Reader scores	
	All readers		Subset of readers*		Range of scores	All readers	Subset of readers*
	Median	Mean (Range)	Median	Mean (Range)		**Mean (Range)	Mean (Range)
BASELINE SCORES							
Inter-reader smICC (total score)	0.83	0.81 (0.57– 0.95)	0.91	0.91 (0.90– 0.94)	**0-21**	3.55 (0.1 – 13.6)	4.04 (0.25 – 14.5)
Achilles tendon (Kappa – lesion wise)							
Peri-tendon hypersignal	0.64	0.62 (0.29–0.87)	0.79	0.78 (0.68– 0.87)	**0-3**	0.47 (0 – 2.00)	0.41 (0 – 2.00)
Intra-tendon hypersignal	0.55	0.51 (0.12– 0.89)	0.79	0.81 (0.77– 0.89)	**0-3**	0.34 (0 – 1.50)	0.33 (0 – 1.50)
Retro-calcaneal bursitis	0.55	0.49 (-0.12-0.93)	0.57	0.60 (0.42– 0.78)	**0-3**	0.34 (0 – 2.1)	0.40 (0 – 2.5)
Bone marrow edema	0.86	0.86 (0.75– 0.97)	0.89	0.88 (0.84– 0.92)	**0-3**	0.38 (0 – 3.00)	0.39 (0 – 3.00)
Plantar fascia (Kappa – lesion wise)							
Peri-aponeurosis hypersignal	0.65	0.60 (0.06– 0.89)	0.80	0.80 (0.66– 0.89)	**0-3**	0.85 (0 – 2.80)	1.01 (0 – 3.00)
Intra-aponeurosis hypersignal	0.55	0.46 (-0.0– 0.93)	0.84	0.84 (0.79– 0.93)	**0-3**	0.56 (0 – 2.00)	0.80 (0 – 3.00)
Bone marrow edema	0.89	0.89 (0.76– 0.98)	0.87	0.88 (0.81– 0.96)	**0-3**	0.61 (0 – 2.90)	0.65 (0 – 3.00)
CHANGE SCORES							
Inter-reader smICC (total score)	0.82	0.80 (0.57– 0.92)	0.84	0.85 (0.79– 0.85)	**0-21**	1.54 (0.1 – 4.9)	1.99 (0 – 6.25)
Achilles tendon (Kappa – lesion wise)							
Peri-tendon hypersignal	0.49	0.47 (0.21– 0.75)	0.50	0.53 (0.42– 0.75)	**0-3**	0.20 (0 – 0.8)	0.17 (0 – 1.00)
Intra-tendon	0.41	0.41	0.54	0.52	**0-3**	0.15	0.17

hypersignal		(0.09– 0.63)		(0.35– 0.63)		(0 – 1.00)	(0 – 1.25)
Retro-calcaneal bursitis	0.24	0.26 (-0.2– 1.00)	0.45	0.42 (0.15– 0.67)	**0-3**	0.21 (0 – 0.8)	0.29 (0 – 1.00)
Bone marrow edema	0.52	0.53 (0.30– 0.82)	0.47	0.54 (0.45– 0.76)	**0-3**	0.14 (0 – 1.1)	0.20 (0 – 1.25)
Plantar fascia (Kappa – lesion wise)							
Peri-aponeurosis hypersignal	0.66	0.61 (0.17– 0.87)	0.77	0.78 (0.70– 0.85)	**0-3**	0.33 (0 – 1.30)	0.42 (0 – 1.75)
Intra-aponeurosis hypersignal	0.53	0.42 (-0.08-0.80)	0.62	0.64 (0.57– 0.77)	**0-3**	0.24 (0 – 0.90)	0.40 (0 – 1.25)
Bone marrow edema	0.78	0.77 (0.57– 0.94)	0.81	0.79 (0.69– 0.88)	**0-3**	0.28 (0 – 1.30)	0.58 (0 – 1.25)

smICCs - single measures intraclass correlation coefficient by two-way random effects, absolute agreement for sum scores (patient level).

* One participating radiologist + three rheumatologist with best individual ICC with other readers for inflammatory pathologies in exercise 2

**Each patient's score was calculated as the mean of all readers. The presented mean and ranges are means/ranges of these values. (Range of readers' mean scores)

Readers: AJM, DG, FG, IE †, MS, MØ ‡, PB, SK ‡, SJP, VF ‡, † Musculoskeletal radiologist. ‡ Three rheumatologists with best individual ICC for inflammatory pathologies in exercise 2.