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

Ashish J. Mathew, Simon Krabbe, Iris Eshed, Frédérique Gandjbakhch, Paul Bird, Susanne J. Pedersen, Maria S. Stoenoiu, Violaine Foltz, Daniel Glinatsi, Robert G. Lambert, Kay Geert A. Hermann, Walter P. Maksymowycz, Ida K. Haugen, Jacob L. Jaremko, René P. Poggenborg, Joel Paschke, Jean-Denis Laredo, Philippe Carron, Philip G. Conaghan, and Mikkel Østergaard

ABSTRACT. Objective. To develop and validate an enthesitis magnetic resonance imaging (MRI) scoring system for spondyloarthritis/psoriatic arthritis, using the heel as model.

Methods. Consensus definitions of key pathologies and 3 heel enthesitis multireader scoring exercises were done, separated by discussion, training, and calibration.

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

Conclusion. The proposed definitions and Heel Enthesis Scoring System (HEMIRS) are reliable among trained readers and promising for clinical trials. (First Release April 2019; J Rheumatol 2019;46:1232–8; doi:10.3899/jrheum.181093)

Key Indexing Terms: ENTHESOPATHY MAGNETIC RESONANCE IMAGING SPONDYLOARTHRPATHY PSORIATIC ARTHRITIS OMERACT
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. Magnetic resonance imaging (MRI) is a sensitive method for detecting enthesitis in peripheral SpA and the only method allowing detection of perienthesal osteitis. 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. Our aim was to create consensus-based MRI definitions of key enthesitis pathologies and through multireader exercises to develop and validate an MRI score for assessing enthesitis in patients with SpA, focusing on the heel region.

MATERIALS AND METHODS
The Outcome Measures in Rheumatology (OMERACT) MRI in Arthritis Working Group initially performed a systematic literature review (SLR) of studies with MRI being used for assessment of enthesitis. Based on this SLR, MRI sequences for optimal visualization of enthesitis were identified, and MRI definitions of key enthesitis pathologies were decided by consensus among group members through meetings/e-mails. The heel region (insertions of Achilles tendon and plantar fascia) was chosen for initial testing because of its frequent involvement. Three multireader exercises, with consensus discussion and calibration in-between, were then performed. A graphical data entry schematic was created, and subsequently a Web-based interface that simultaneously displayed DICOM (Digital Imaging and Communications in Medicine) images and the data entry schematic (Figure 1). 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 MRI [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 MRI (sagittal T2wFS only) of 21 patients with SpA 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 was analyzed separately between the participating radiologists and the 3 rheumatologists with best overall intraclass correlation coefficient (ICC) for inflammatory pathologies in exercise 2.

**Figure 1.** Line drawing of the scoring sheet used in the scoring exercises. BME: bone marrow edema.
**RESULTS**

**Definitions of key pathologies.** Key enthesal 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, and published OMERACT MRI definitions for comparable conditions. The selected pathologies were intratendon hypersignal (enthesal tendonitis), peritendon hypersignal (enthesal peritendonitis), bone marrow edema (BME; enthesal osteitis), bursitis, tendon thickening, enthesophyte, enthesal bone erosion, and intratendon 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 (Figure 3). A T1W sequence prior to contrast injection (T1-pre-Gd) was considered helpful in determining the exact localization of inflammatory pathologies because of 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 (rheumatoid arthritis magnetic resonance imaging) and PsAMRIS (psoriatic arthritis magnetic resonance imaging) systems, and to create a total enthesal inflammation score by summation of scores of all inflammatory variables (intratendon hypersignal on T2w/STIR sequences, peritendon hypersignal, BME, and bursitis). Similarly, a total enthesal structural damage score by summation of structural scores (enthesophyte, bone erosion, tendon thickening) was developed. Intratendon 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 predefined pathologies. Exercise 1 was used for initial learning, calibration, and identification of pitfalls. Mean pairwise interreader single-measure ICC 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 to very good for various lesion types and their sum scores (Table 2). When limiting the analyses to 3 participating musculoskeletal radiologists and 3 rheumatologists with best ICC for inflammatory pathologies in exercise 2, reliability improved to moderate to very good. For this subset of readers, median single-measure ICC for total inflammation scores was 0.85, while for total structural damage scores was 0.68. Median κ for different inflammatory pathologies varied from 0.60 to 0.89, and for individual structural pathologies from 0.41 to 0.78. Average-measure ICC based on 2 readers among the preselected 6 readers (median 0.92 for total inflammatory score, 0.81 for total damage scores) were better than the single-measure ICC.

Exercise 3. This exercise included 2 timepoint images, in which inflammatory pathologies were scored. Mean pairwise interreader ICC and lesion-wise κ agreement demonstrated moderate to good reliability when all readers were considered (Table 3). The subset of readers (3 rheumatologists with best agreement for inflammatory measures 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 ICC for total inflammation was 0.91, while it was 0.84 for change in score. Median average-measure ICC based on 2 readers [status: median 0.97 (range 0.95–0.97), change: 0.92 (0.89–0.96)] were higher than single-measure ICC. Using 3 readers demonstrated numerically higher average-measure ICC [status: median 0.97 (0.97–0.97), change 0.94 (0.94–0.95)].

The Heel Enthesitis Scoring System (HEMIRIS) showed moderate responsiveness, with SRM of 0.70 (95% CI 0.38–1.05) for all readers in exercise 3.

**DISCUSSION**

To the best of our knowledge, our study is the first international consensus effort toward development of a comprehensive MRI scoring system, combined with MRI definitions and reader rules, for enthesitis in patients with SpA. The work was informed by an SLR8, which clarified knowledge gaps and the need for development of a validated MRI enthesitis scoring system to be used as outcome measure in clinical trials. Enthesitis, often located at the heels, is a typical feature of SpA and is easily accessible for MRI12. Further, enthesitis in SpA may show changes both in inflammation (such as BME and perienthesial 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 multirater scoring exercises focused on the heel region, using an intuitive Web-based data entry and image display platform. The preliminary OMERACT-HEMIRIS showed good interreader agreement for status scores and for change over time in inflammatory measures. Considering that baseline heel enthesitis was not mandatory in exercise 3, the moderate SRM (0.70) supports that responsiveness of the HEMIRIS score would likely be good in trials with baseline enthesitis as an inclusion criterion. Thus, HEMIRIS appears promising for further validation and future use in randomized controlled trials.

---

Table 1. Magnetic resonance imaging definitions of key enthesial pathologies.

| Pathology                          | Definition                                                                 |
|-----------------------------------|---------------------------------------------------------------------------|
| 1. Intratendon hypersignal (entheseal tendonitis) | Signal characteristics consistent with increased water content/inflammation* within the tendon/ligament/aponeurosis close to its insertion |
| 2. Peritendon hypersignal (entheseal peritendonitis) | 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 marinated 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 (T1W) 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 short-tau inversion recovery/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. T2wFS: T2w fat-suppressed [images].
The strengths of this initiative include taking an SLR as starting point to clarify unmet need, the involvement of experienced MRI researchers in the development of consensus-based definitions and scoring systems, and 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 subanalysis 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 variables. 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 2 readers (compared to 1) is used in the final study analysis; the average-measure ICC for 2 readers were markedly higher (0.92–0.95 for inflammation

**MRI sequences:** STIR/T2wFS or, alternatively T1w post-Gd; T1w without contrast (not mandatory if only inflammation is assessed)

**Imaging planes:** Achilles tendon: Sagittal and preferably axial; Plantar aponeurosis: Sagittal and preferably coronal

**Area to score:** At the heel region, the entheses are evaluated within 1 cm from the tendon/aponeurosis insertion.

**Scoring procedure:**

**Enthesal soft tissue inflammation**

- If T1w post-Gd images are available, enthesal soft tissues are assessed on these and the intratendon/peritendon/bursal hypersignal is defined as above-normal post-gadolinium enhancement on T1w images
- If only STIR/T2wFS images are available, enthesal soft tissues are assessed on these and the intratendon/peritendon/bursal hypersignal is defined as high signal intensity on STIR/T2wFS images
- Grading scale is 0-3 based on thirds of the maximum potential volume of enhancing soft tissue:
  
  - Score 0 – normal; 1 – mild; 2 – moderate and 3 – severe.

**Enthesal osteitis**

- If STIR/T2wFS images are available, enthesal osteitis is assessed on these defined as a lesion within the enthesal bone marrow with ill-defined margins and high signal intensity on STIR/T2wFS images (“bone marrow edema”)
- If only T1w-post Gd images are available, enthesal osteitis is assessed on these, and defined as a lesion within the enthesal bone marrow, with ill-defined margins, which shows above-normal enhancement on T1w-post-Gd images (“bone marrow post-contrast enhancement”)
- Grading 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 the bone is edematous; 2: 34-66% of the bone is edematous; 3: 67-100% of the bone is edematous.

**Enthesal structural damage variables**

- Enthesal structural damage variables are scored based on T1w pre-Gd images.

*Figure 3.* Reader rules for heel enthesitis. STIR: short-tau inversion recovery; T2wFS: T2W fat-suppressed; T1W: T1-weighted; Gd: gadolinium.
### Table 2. Exercise 2: Single-measure interreader intraclass correlation coefficient (ICC; sum scores), quadratic weighted $\kappa$ (individual component scores, per lesion), and mean scores of all readers.

| Variables                        | Reproducibility, smICC and $\kappa$ | Reader Scores |
|----------------------------------|---------------------------------------|---------------|
|                                  | All Readers Median | All Readers Mean (Range) | Subset of Readers Median | Subset of Readers Mean (Range) | Range of Scores **Mean (Range)** | Range of Scores **Mean (Range)** |
| **Inflammatory pathologies**     |                         |                 |                         |                              |                         |                           |
| Interreader 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)          |
| Interreader quadratic weighted $\kappa$ (lesions level) |                         |                 |                         |                              |                         |                           |
| Achilles tendon                  |                         |                 |                         |                              |                         |                           |
| Peritendon 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)             |
| Intratendon 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)             |
| Retrocalcaneal bursitis          | 0.47                    | 0.45 (–0.06 to 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                   |                         |                 |                         |                              |                         |                           |
| Perianeural 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)             |
| Perianeural hypertenderness       | 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 (0.11–0.98) | 0.86                    | 0.86 (0.73–0.94)             | 0–3                      | 0.66 (0–2.47)             | 0.74 (0–2.67)             |
| **Structural pathologies**       |                         |                 |                         |                              |                         |                           |
| Interreader smICC (patient level)|                          |                 |                         |                              |                         |                           |
| Total structural damage score    | 0.27                    | 0.35 (–0.04 to 0.85) | 0.68 | 0.66 (0.37–0.85) | 0–18 | 1.54 (0.2–4.4) | 2.33 (0.33–7.00) |
| Interreader quadratic weighted $\kappa$ (lesion level) |                         |                 |                         |                              |                         |                           |
| Achilles tendon                  |                         |                 |                         |                              |                         |                           |
| Tendon thickness                 | 0.52                    | 0.48 (0–0.92)   | 0.76                    | 0.72 (0.4–0.92)              | 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 to 1.0) | 0.41 | 0.37 (0–0.87) | 0–3 | 0.13 (0–0.87) | 0.22 (0–1.33) |
| Intratendon hypersignal on T1W@  | 0.30                    | 0.33 (–0.09 to 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 to 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 to 0.64) | 0.00 | 0.03 (–0.05 to 0.14) | 0–3 | 0.06 (0–0.27) | 0.11 (0–0.5) |
| Bone spur                        | 0.00                    | 0.12 (–0.18 to 0.76) | 0.42 | 0.4 (–0.18 to 0.76) | 0–3 | 0.17 (0–0.53) | 0.28 (0–1.17) |
| Intratendon hypersignal on T1W@  | 0.21                    | 0.25 (–0.19 to 0.84) | 0.40 | 0.42 (0.05–0.83) | 0–3 | 0.13 (0–1.27) | 0.49 (0–2.00) |

* Three participating radiologists and 3 rheumatologists 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 present mean and ranges are means/ranges of these values (range of readers’ mean scores). Readers: AJM, DG, FG, HH, IE†, KGH†, MS, M0†, PB, RGL†, SK†, SIP, VF†, WM (exercises 1 and 2); JI† (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. smICC: single-measures intraclass correlation coefficient by 2-way random effects, absolute agreement for sum scores (patient level); T1W: T1-weighted.

total status/change score in last exercise) than single-measure ICC. This will be relevant in real-life clinical trials where 2 independent readers generally score images.

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

ACKNOWLEDGMENT

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

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. Polacheck A, Li S, Chandran V, Gladman DD. Clinical enthesis 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), Nails 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 Enthesitis 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, Alothoff CE, Ashbach P, et al. MRI of enthesis of the appendicular skeleton in spondyloarthropathy. Ann Rheum Dis 2007;66:1553-9.
Table 3. Exercise 3: Baseline and change single-measure interreader intraclass correlation coefficient (ICC; sum scores), quadratic weighted \( \kappa \) (individual component scores, per lesion), and mean scores of all readers.

| Variables                        | Reproducibility, smICC and \( \kappa \) | Range of Scores | Reader Scores |
|----------------------------------|------------------------------------------|-----------------|--------------|
|                                  | All Readers | Median (Mean (Range)) | Subset of Readers* | Median (Mean (Range)) | All Readers | Median (Mean (Range)) | Subset of Readers* | Median (Mean (Range)) |
| Baseline scores                  |             |                         |                   |                         |             |                     |                   |                     |
| Interreader smICC (total score)  |             | 0.83 (0.57–0.95)        | 0.91 (0.90–0.94)  | 0–21                     | 3.55 (0.1–13.6) | 4.04 (0.25–14.5)    |                   |                     |
| Achilles tendon (lesion)          |             | 0.64 (0.29–0.87)        | 0.79 (0.68–0.87)  | 0–3                      | 0.47 (0–2.00)  | 0.41 (0–2.00)       |                   |                     |
| Periarticular hypersignal         |             | 0.55 (0.12–0.89)        | 0.79 (0.81–0.77)  | 0–3                      | 0.34 (0–1.50)  | 0.33 (0–1.50)       |                   |                     |
| Radiological mastitis             |             | 0.55 (0.49–1.02 to 0.93) | 0.57 (0.60–0.42)   | 0–3                      | 0.34 (0–2.1)   | 0.40 (0–2.5)        |                   |                     |
| Bone marrow edema                |             | 0.86 (0.75–0.97)        | 0.89 (0.88–0.84)  | 0–3                      | 0.38 (0–3.00)  | 0.39 (0–3.00)       |                   |                     |
| Change scores                    |             | 0.82 (0.57–0.92)        | 0.84 (0.85–0.79)  | 0–21                     | 1.54 (0–4.9)   | 1.99 (0–6.25)       |                   |                     |
| Achilles tendon (lesion)          |             | 0.49 (0.21–0.75)        | 0.50 (0.53–0.42)  | 0–3                      | 0.20 (0–0.8)   | 0.17 (0–1.00)       |                   |                     |
| Periarticular hypersignal         |             | 0.41 (0.09–0.63)        | 0.54 (0.52–0.35)  | 0–3                      | 0.15 (0–1.00)  | 0.17 (0–1.25)       |                   |                     |
| Radiological mastitis             |             | 0.24 (0.02 to 0.100)    | 0.45 (0.42–0.15)  | 0–3                      | 0.21 (0–0.8)   | 0.29 (0–1.00)       |                   |                     |
| Bone marrow edema                |             | 0.52 (0.30–0.82)        | 0.47 (0.54–0.45)  | 0–3                      | 0.14 (0–1.1)   | 0.20 (0–1.25)       |                   |                     |
| Plantar fascia (lesion)           |             | 0.66 (0.17–0.87)        | 0.77 (0.70–0.85)  | 0–3                      | 0.33 (0–1.30)  | 0.42 (0–1.75)       |                   |                     |
| Periarticular hypersignal         |             | 0.53 (0.42–0.08 to 0.80) | 0.62 (0.64–0.57)  | 0–3                      | 0.24 (0–0.90)  | 0.40 (0–1.25)       |                   |                     |
| Bone marrow edema                |             | 0.78 (0.57–0.94)        | 0.81 (0.79–0.69)  | 0–3                      | 0.28 (0–1.30)  | 0.58 (0–1.25)       |                   |                     |

*One participating radiologist and 3 rheumatologists 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, DFG, HE1, MS, MS2, PB, SK2, SJP, VF. †Musculoskeletal radiologist. ‡Three rheumatologists with best individual ICC for inflammatory pathologies in exercise 2. smICC: single-measures intraclass correlation coefficient by 2-way random effects, absolute agreement for sum scores (patient level).

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 enthethesopathies may not present as focal insertional disorders. Arthritis Rheum 2004;50:3306-13.

7. Dougados M1, Combe B, Braun J, Landewe R, Sibilia J, Cantagrel A, et al. A randomised, multicentre, double-blind, placebo-controlled trial of etanercept in adults with refractory heel enthesis in spondyloarthropathies: the HEEL trial. Ann Rheum Dis 2010;69:1430-35.

8. Mathew AJ, Krabbe S, Kirubakaran R, Barr AJ, Conaghan PG, Bird P, et al. Utility of magnetic resonance imaging in diagnosis and monitoring enthesis in patients with spondyloarthropathy: an OMERACT systematic literature review. J Rheumatol 2019;46:1207-14.

9. 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.

10. Østergaard M, Peterfy CG, Bird P, Gandjbakhch F, Glinatsi D, Eshed I, et al. The OMERACT RA in arthritis (OMERACT RA-AR) MRI scoring system. J Rheumatol 2019;46:1385-6.

11. Østergaard M, Peterfy CG, Bird P, Gandjbakhch F, Glinatsi D, Eshed I, et al. The OMERACT RA in arthritis (OMERACT RA-AR) MRI scoring system: definitions of key pathologies, suggested MRI sequences, and preliminary scoring system for PsA hands. J Rheumatol 2009;36:1816-24.

12. Østergaard M, Peterfy CG, Bird P, Gandjbakhch F, Glinatsi D, Eshed I, et al. The OMERACT RA in arthritis (OMERACT RA-AR) MRI scoring system: definitions of key pathologies, suggested MRI sequences, and preliminary scoring system for PsA hands. J Rheumatol 2009;36:1816-24.