Seminars in Arthritis and Rheumatism 51 (2021) 933–939

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Joint and entheseal inflammation in the knee region in spondyloarthritis - reliability and responsiveness of two OMERACT whole-body MRI scores

Marie Wetterslev[a,b,*], Walter P Maksymowych[c,d], Robert GW Lambert[e,f], Iris Eshed[g], Susanne J Pedersen[a], Maria S Stoenoiu[h], Simon Krabbe[a], Paul Bird[i], Violaine Foltz[j], Ashish J Mathew[a,b], Frédérique Gandjbakhch[j], Joel Paschke[d], Philippe Carron[k,l], Gabriele De Marco[m,n], Helena Marzo-Ortega[m,n], Anna EF Poulsen[a], Jacob L Jaremko[e], Philip G Conaghan[m,n], Mikkel Østergaard[a,b]

[a] Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Copenhagen, Denmark
[b] Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
[c] Department of Medicine, University of Alberta, Edmonton, Alberta, Canada
[d] CARE Arthritis, Edmonton, Alberta, Canada
[e] Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, Alberta, Canada
[f] Medical Imaging Consultants, Edmonton, Alberta, Canada
[g] Department of Diagnostic Imaging, Sheba Medical Center, Affiliated to the Sackler School of Medicine, Tel-Aviv University, Israel
[h] Department of Rheumatology, Cliniques Universitaires Saint-Luc, Institut de Recherche Expérimentale et Clinique, Université catholique de Louvain, Bruxelles, Belgium
[i] Division of Medicine, University of New South Wales, Sydney, Australia
[j] Department of Rheumatology, Sorbonne University, APHP, Pitié-Salpêtrière Hospital, Paris, France
[k] Department of Rheumatology, Ghent University Hospital, Ghent, Belgium
[l] VIB Inflammation Research Centre, Ghent University, Ghent, Belgium
[m] Leeds Institute of Rheumatic and Musculoskeletal medicine, University of Leeds, Leeds, UK
[n] Leeds Teaching Hospitals NHS Trust, NIHR Leeds Biomedical Research Centre, Leeds, UK

ARTICLE INFO

Keywords:
OMERACT
Spondyloarthritis
Knee
Whole-body MRI
MRI-WIPE
KIMRISS

ABSTRACT

Objective: To perform region-based development of whole-body MRI through validation of knee region scoring systems in spondyloarthritis (SpA).
Methods: Assessment of knee inflammatory pathologies using 2 systems, OMERACT MRI Whole-body score for Inflammation in Peripheral joints and Entheses (MRI-WIPE) and Knee Inflammation MRI Scoring System (KIMRISS), in 4 iterative multi-reader exercises.
Results: In the final exercise, reliability was mostly good for readers with highest agreement in previous exercise. Median pairwise single-measure ICCs for osteitis and synovitis/effusion status/change were 0.71/0.48 (WIPE-osteitis), 0.48/0.77 (WIPE-synovitis/effusion), 0.59/0.91 (KIMRISS-osteitis) and 0.92/0.97 (KIMRISS-synovitis/effusion). SRMs were 0.74 (WIPE-synovitis/effusion) and 0.78 (KIMRISS-synovitis/effusion).
Conclusion: MRI-WIPE and KIMRISS may both be useful in SpA whole-body evaluation studies.

Introduction

Inflammation in peripheral joints and entheses is common in spondyloarthritis (SpA) including psoriatic arthritis (PsA) [1,2]. Magnetic resonance imaging (MRI) allows detailed assessment of inflammation in both soft tissue and bone [3,4], traditionally in a limited anatomical area. Whole-body MRI (WB-MRI) allows assessment of the overall inflammatory status of arthritis patients, including

Abbreviations: axSpA, axial spondyloarthritis; HEMRIS, OMERACT Heel enthesitis MRI scoring system; ICC, intraclass correlation coefficient; kappa, Cohen's kappa; KIMRISS, Knee inflammation MRI scoring system; MR, magnetic resonance; MRI, magnetic resonance imaging; MRI-WIPE, OMERACT MRI whole-body score for inflammation in peripheral joints and entheses in inflammatory arthritis; OA, osteoarthritis; OMERACT, outcome measures in rheumatology; PsA, psoriatic arthritis; PSAMRIS, psoriatic arthritis MRI scoring system; RETIC, real-time iterative calibration; spondyloarthritis, SpA; SRM, standardized response mean; T1w post-Gd, T1-weighted post-gadolinium; T2wFS, T2-weighted fat-suppressed; STIR, short-tau inversion recovery; WB-MRI, whole-body MRI

* Corresponding author at: Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Rigshospitalet, Glostrup, Valdemar Hansen Vej 17, 2600 Glostrup, Denmark.
E-mail address: marie.wetterslev@regionh.dk (M. Wetterslev).

https://doi.org/10.1016/j.semarthrit.2021.05.013
0049-0172/

joints and entheses [5,6], and is therefore well suited for assessment of spondyloarthritides. The Outcome Measures in Rheumatology (OMERACT) MRI Whole-body score for Inflammation in Peripheral joints and Entheses (MRI-WIPE) has been developed and preliminarily validated (construct validity, reproducibility and responsiveness) for the entire body, including the knee, but not separately for the knee joint region [7–9]. Detailed MRI scoring systems exist for heels (OMERACT Heel Enthesitis MRI Scoring System (HEMRIS)) [10], hands and feet (OMERACT Psoriatic Arthritis MRI Scoring System (PsAMRIS)) [11], but although knee arthritis is a key cause of functional impairment, no detailed MRI scoring system for knee inflammation in SpA has been published. In 2019, the international OMERACT MRI in Arthritis Working Group decided to further develop and validate WB-MRI in SpA by investigating methods, including MRI-WIPE, with a modular, i.e. region-based, approach. The Knee Inflammation MRI Scoring System (KIMRISS) is a granular (finely detailed) semiquantitative scoring system developed and validated in patients with osteoarthritis (OA), in whom it showed good reliability for status and change in bone marrow lesions [12].

We therefore aimed to investigate and compare MRI-WIPE and KIMRISS for assessment of inflammation, i.e. osteitis (bone marrow edema), synovitis and soft tissue inflammation in the knee region of patients with SpA and evaluate interreader agreement for status and change, responsiveness and correlation between the two methods.

Materials and method

Materials

From January to September 2020 radiologists and rheumatologists from 7 countries participated in 4 web-based multi-reader exercises applying MRI-WIPE and KIMRISS and 6 web-conferences. An online real-time iterative calibration (RETIC) module for KIMRISS was available [13] along with online instructional presentations for MRI-WIPE.

In all exercises, anonymized whole-body MR knee images (i.e. images obtained as part of a WB-MRI-examination, with lower resolution than conventional dedicated MRI), blinded for chronology, were uploaded to a web-based interface hosted securely by CARE Arthritis, Edmonton, Canada, which displayed the images and data entry schematics for WIPE-knee as well as superimposing interactive overlays for evaluation using KIMRISS. Images were scored according to the semiquantitative OMERACT MRI-WIPE system [7] and the more detailed knee MRI scoring method KIMRISS [12] (Appendix) to validate the scoring systems in accordance with the OMERACT Filter (2.1) Instrument Selection Algorithm (OFISA) [14]. Images were assessed independently by readers with varying expertise in MRI and in the two scoring methods.

In exercise 1, performed to train inexperienced readers and identify pitfalls, sagittal T1-weighted (T1w) and short-tau inversion recovery (STIR) knee images from 3 cases (axial SpA (axSpA)) were evaluated by 12 readers (2/10 radiologists/rheumatologists).

In exercise 2, sagittal T1w and STIR knee images from 7 cases (axSpA) were evaluated by 9 readers (1/8 radiologist/rheumatologists). Subsequently, difficulties and discrepancies were discussed online to improve consensus and the KIMRISS method for measuring synovitis/effusion was discussed and adapted to SpA. This adapted consensus-based approach was applied in exercise 3 and 4 [15].

In exercise 3, sagittal T1w and STIR knee images of 11 cases with 2 timepoints (axSpA and PsA, 9 of 11 patients before and after initiation of tumor necrosis factor (TNF) inhibitor) were evaluated by 9 readers (2/7 radiologists/rheumatologists). For assessing the reliability among more calibrated and experienced readers agreement was analysed for the 3 readers with the overall highest interreader agreement. Subsequently, selected reference images for WIPE-knee were discussed to obtain consensus.

In exercise 4, STIR/T2-weighted fat-suppressed (T2wFS) knee images from 10 cases with timepoints before and after TNF inhibitor (axSpA and peripheral SpA) and sagittal T1w post-gadolinium (post-Gd) images reconstructed from axial Dixon sequences from 10 cases with one timepoint (PsA) were evaluated by 9 readers (2/7 radiologists/rheumatologists). In all exercises, patients did not necessarily have any baseline inflammation in the knee region.

Statistics

For exercises 3 and 4, agreement at lesion level (only possible for MRI-WIPE) was assessed using Cohen's kappa (kappa), quadratically weighted [16] and agreement at patient level was assessed using single measure intraclass correlation coefficient (ICC), two way mixed model and absolute agreement definition [17,18]. The correlation between WIPE-knee and KIMRISS was assessed using Spearman's rho. Wilcoxon signed-rank test and standardized response mean (SRM) [19] were applied to evaluate changes between timepoints. Statistical analyses were made in SPSS version 25.0 or R version 3.6.1; $p<0.05$ was considered statistically significant.

Results

Physicians from 7 countries participated in exercises and web-meetings. Six out of the 12 readers who participated in exercises completed the KIMRISS calibration modules before exercise 3. The readers with overall highest interreader agreement in exercise 3 did not finalize calibration modules prior to exercise 3 but were experienced readers (1/2 radiologist/rheumatologist) and developers of the scoring methods.

A modification of the KIMRISS synovitis/effusion assessment was developed, discussed and finalized before exercise 3. Using this, synovitis/effusion was measured in a predefined area (Fig. 1) on consecutive slices of fluid-sensitive sequences as the largest diameter perpendicular to the longest axis of the largest focus of synovial fluid/thickening.

In exercise 3, interreader reliability between reader pairs for status and change in sum scores varied from poor to good for both methods (Table 1). Large variation was seen between reader pairs and mean pairwise interreader single-measure ICCs and kappas improved markedly when looking at mean ICCs from the 3 readers with overall highest interreader ICCs in exercise 3.

In exercise 4, agreement (ICC and Kappa) between reader pairs also improved markedly when looking at the 3 readers with highest interreader agreement in exercise 3. Single-measure ICCs varied from poor to very good for osteitis and synovitis/effusion for status and change and were 0.71/0.48 (WIPE-knee osteitis), 0.48/0.77 (WIPE-knee synovitis/effusion), 0.59/0.91 (KIMRISS osteitis) and 0.92/0.97 (KIMRISS synovitis/effusion) (Table 1). ICCs were most often numerically higher for KIMRISS.

Regarding responsiveness the Wilcoxon signed-rank test in exercise 4 showed a significant change between timepoints in synovitis/effusion for both WIPE-knee and KIMRISS (Table 2). SRM for synovitis/effusion was moderate for both MRI-WIPE and KIMRISS (0.74 and 0.78) and lower for osteitis. The two methods correlated significantly regarding status for osteitis and synovitis/effusion and for change in synovitis/effusion.

Discussion

In this OMERACT study a modular approach to WB-MRI was applied. Inflammation in the knee region was assessed in patients with SpA using the 2 different scoring methods, MRI-WIPE and KIMRISS. The study showed mostly good agreement for status and change in osteitis and good to very good agreement for status and change in synovitis/effusion, numerically highest for KIMRISS. The KIMRISS

Measuring knee synovitis/effusion according to KIMRISS

MRI sequences: STIR/T2wFS or alternatively T1w post-Gd
Imaging plane: Sagittal
Definition of area:
"Medial area":
Areas medial to the patella (patella not visible) and anterior to the mid-coronal plane of the tibial plateau.
"Lateral area":
Areas lateral to the patella (patella not visible) and anterior to mid-coronal plane of tibial plateau.
"Anterior area":
Areas superior, inferior or posterior to the patella (patella visible) (includes supra-, retro-, infra-patellar areas) and anterior to the mid-coronal plane of the tibial plateau.
"Posterior area":
Areas posterior to the mid-coronal plane of the tibial plateau (including Baker's cyst).
Measuring procedure:

- Knee synovitis/effusion is evaluated on every slice where synovitis/effusion is present.
- In each slice look at all 4 areas ("medial area", "lateral area", "anterior area", "posterior area") and find the area with the most synovitis/effusion. Only one area per slice is measured.
- The area with the most synovitis/effusion is defined as the largest diameter perpendicular to the longest axis.
- In the area with most synovitis/effusion the largest perpendicular diameter is measured (see examples).
- When the largest diameter on a slice has been measured, choose the area by ticking the appropriate box. This is done on every slice where synovitis/effusion is measured.

Score per slice:
The measured perpendicular diameter determines the score.
0: 0-1.9 mm
1: 2-4.9 mm
2: 5-9.9 mm
3: 10-19.9 mm
4: ≥ 20 mm

Fig. 1. KIMRISS reader rules for assessment of knee synovitis/effusion. KIMRISS, Knee Inflammation MRI Scoring System; STIR, short-tau inversion recovery; T2wFS, T2-weighted fat-supressed; T1w post-Gd, T1-weighted post-gadolinium.

method for assessment of synovitis/effusion was further developed and improved through consensus-based discussions in the OMERACT MRI in Arthritis Working Group.

This is the first study of KIMRISS outside OA. Also, this is the first study where the OMERACT MRI-WIPE is used to evaluate individual regions on WB-MRI, i.e. a modular approach. Large variation was seen between reader pairs, but the methods seemed reliable and sensitive to change among experienced and well-calibrated readers.

It should be noticed that WIPE-knee and KIMRISS measure 2 different things. KIMRISS provides granular assessment of osteitis and synovitis/effusion in the knee joint itself, which may contribute to a higher reproducibility and responsiveness, and does not consider entheseal regions. In contrast WIPE-knee provides a less granular assessment of soft tissue and bone marrow inflammation at the knee joint, but adds assessment of entheses such as quadriceps and patella tendon insertions. Since enthesitis is an important domain in PsA/SpA, its assessment is an advantage of WIPE. Thus, the methods cannot be directly compared and are complementary rather than competing.

Limitations in our study include the relatively low number of cases. Moreover, the cases varied regarding follow-up time and did not necessarily have inflammation in the knee region. This may influence the responsiveness of the measures. Especially in exercise 4 the observed range of scores for osteitis was low compared to the maximum possible score and only minimal change over time was seen. This would tend to cause lower ICCs. It would have been optimal to have images with more synovitis/effusion, osteitis and change over time. However, the image material available was limited and did not allow us to choose an optimal sample collection. STIR images were used in the first exercises while T2wFS and T1w post-Gd images were also used in exercise 4 potentially influencing reader agreement, particularly for less experienced readers. Furthermore, there was a large variation in experience of readers and not all completed the KIMRISS calibration modules available. This was not considered mandatory, since the study was preliminary. Going forward and in order to optimize reader performance, a reference atlas for WIPE-knee and obligatory completion of prespecified calibration exercises should be included in future developments.

To summarize, two complementary semiquantitative MRI scoring systems, MRI-WIPE and KIMRISS, allow assessment of knee-region inflammation in patients with SpA including PsA. The methods showed mostly good to very good agreement between reader pairs and acceptable sensitivity to change. Our results imply that careful attention to reader calibration is necessary to optimize performance.

In conclusion, assessment of inflammation in the knee region is an important part of WB-MRI interpretation in spondyloarthritis. WIPE-knee and KIMRISS are promising tools for further validation and use in randomized controlled trials in SpA including PsA.

Declaration of Competing Interest

WPM is Chief Medical Officer CARE Arthritis Limited and has acted as a paid consultant/participated in advisory boards for AbbVie, Boehringer Ingelheim, Celgene, Eli Lilly, Galapagos, Janssen, Novartis, Pfizer and UCB; received research and/or educational grants from AbbVie, Novartis, Pfizer and UCB; received speaker fees from AbbVie, Janssen, Novartis, Pfizer and UCB. RGWL has received consulting fees from CARE Arthritis, Parexel and Pfizer. SJP has been an advisory board member for AbbVie and Novartis; received research support from AbbVie, MSD, and Novartis; received speaker fees from MSD, Pfizer, AbbVie, Novartis and UCB. PB participated in advisory boards and received speaker fees from Janssen, Abbvie, UCB, Celgene, BMS, Novartis, Pfizer, Gilead, Eli-Lilly. PC has received research grants from

Table 1
MRI-WIPE knee and KIMRISS interreader reliability for exercises 3 and 4.

			MRI-WIPE Knee						KIMRISS			
			Osteitis			Synovitis/effusion			Osteitis		Synovitis/effusion	
	No. patients (cases)	Type of score	Mean score	ICC	Kappa	Mean score	ICC	Kappa	Mean score	ICC	Mean score	ICC
Exercise 3 (9 readers)	11	Status	3.6 (0–16)	0.57 (−0.06–0.98)	0.39 (0.04–0.74)	1.8 (0–4)	0.47 (0.05–0.85)	0.44 (0.03–0.77)	32.3 (1–224)	0.87 (0.66–0.99)	29.9 (11–60)	0.34 (−0.62–0.87)
	11	Change	1.1 (−2–6)	0.53 (0.03–0.90)	0.26 (−0.08–0.50)	0 (−2–1)	0.32 (−0.13–0.76)	0.16 (−0.13–0.64)	27.7 (−9–131)	0.58 (−0.30–0.96)	−1.6 (−33–11)	0.48 (−0.32–0.95)
Exercise 3 (3 readers)	11	Status	3.1 (0–16)	0.83 (0.71–0.97)	0.65 (0.55–0.74)	2.5 (0–5)	0.59 (0.51–0.71)	0.57 (0.48–0.68)	34.4 (0–233)	0.89 (0.83–0.99)	36.5 (16–78)	0.59 (0.08–0.86)
	11	Change	0.9 (−3–6)	0.72 (0.57–0.83)	0.38 (0.33–0.40)	0 (−2–1)	0.63 (0.49–0.76)	0.53 (0.41–0.64)	19.3 (−23–86)	0.46 (0.18–0.83)	−1.8 (−45–17)	0.89 (0.82–0.95)
Exercise 4 (9 readers)	10 (1–10)	Status	3.5 (0–7)	0.62 (−0.01–0.87)	0.47 (0.06–0.76)	2.0 (0–4)	0.44 (0.21–0.79)	0.48 (0.17–0.83)	14.0 (0–29)	0.56 (0.07–0.94)	63.5 (1–122)	0.56 (0.01–0.97)
	10 (11–20)	Status	2.3 (0–7)	0.44 (−0.20–0.93)	0.34 (0.12–0.58)	2.3 (1–4)	0.41 (−0.03–0.83)	0.54 (0.25–0.82	16.3 (1–66)	0.32 (−0.14–0.92)	47.7 (25–76)	0.51 (−0.02–0.98)
	10 (11–20)	Change	−0.25 (−4–5)	0.38 (−0.35–0.94)	0.15 (−0.01–0.76)	−1.0 (−3–1)	0.30 (−0.43–0.89)	0.43 (0.08–0.90)	−0.45 (−37–65)	0.26 (−0.86–0.97)	−14.7 (−48–0.20)	0.48 (−0.39–0.99)
	20 (1–20)	Status	2.9 (0–7)	0.50 (−0.01–0.84)	0.42 (0.25–0.64)	2.1 (0–4)	0.44 (−0.21–0.79)	0.52 (0.33–0.68)	15.2 (0–66)	0.35 (−0.04–0.89)	55.6 (1–122)	0.54 (0.01–0.96)
Exercise 4 (3 readers)	10 (1–10)	Status	3.1 (0–6)	0.80 (0.68–0.87)	0.70 (0.68–0.71)	2.5 (0–5)	0.42 (0.35–0.56)	0.41 (0.31–0.54)	11.7 (0–29)	0.54 (0.38–0.85)	82.8 (1–153)	0.90 (0.86–0.93)
	10 (11–20)	Status	1.5 (0–6)	0.62 (0.43–0.74)	0.37 (0.30–0.48)	2.8 (0–5)	0.58 (0.32–0.83)	0.55 (0.33–0.80)	11.0 (0–36)	0.63 (0.45–0.89)	55.9 (29–93)	0.90 (0.85–0.98)
	10 (11–20)	Change	0.2 (−2–6)	0.48 (0.16–0.66)	0.33 (0.24–0.48)	−1.4 (−5–0)	0.77 (0.70–0.82)	0.76 (0.69–0.85)	5.8 (−27–111)	0.92 (0.90–0.94)	−20.7 (−65–28)	0.97 (0.96–0.98)
	20 (1–20)	Status	2.3 (0–6)	0.71 (0.60–0.80)	0.59 (0.54–0.64)	2.7 (0–5)	0.48 (0.42–0.57)	0.49 (0.43–0.57)	11.4 (0–36)	0.59 (0.39–0.71)	69.4 (1–153)	0.91 (0.87–0.93)

Sum scores are mean (range) of the patients scores (each patient's score is the average of the scores assigned to that patient). MRI-WIPE knee range for osteitis is 0–60 and for synovitis/effusion 0–6 [7]. KIMRISS osteitis total range is 0–500 and range for synovitis/effusion is 0–100 [12,13]. ICC values are mean (range). ICC is 2-way mixed model, single measure, by absolute agreement. ICC values ≤0.49 were considered as poor, 0.50–0.79 as good, ≥0.80 as very good reliability. Kappa values are mean (range). Scorings at lesion level were assessed using Cohen's kappa, quadratically weighted. Kappa 0–0.20 was considered as no agreement, 0.21–0.39 as slight, 0.40–0.59 as weak, 0.60–0.79 as moderate, 0.80–0.90 as strong and >0.90 as almost perfect agreement [16]. Readers: IE[+], MW, MØ*, PB, SJP, WPM* (all exercises), RGWL[+]*, VF (exercise 1, 3, 4), MSS (exercise 1, 2, 4), AJM (exercise 1–3), SK (exercise 1, 2), FG (exercise 1).[+]Musculoskeletal radiologist. *the readers with overall highest agreement in Exercise 3 (MØ, RGWL, WPM).
ICC, intraclass correlation coefficient; Kappa, Cohen's Kappa, quadratic weighted; KIMRISS, Knee Inflammation MRI Scoring System; MRI-WIPE, OMERACT MRI Whole-body score for Inflammation in Peripheral joints and Entheses in inflammatory arthritis.

Table 2
Sensitivity to change and correlation between methods in exercises 3 and 4[1].

Exercise 3		Baseline	Follow-up	Change	p-value	SRM
MRI-WIPE knee						
	Osteitis	3.1 (4.96)	4.0 (4.77)	0.9 (2.44)	0.089	0.37
	Synovitis/effusion	2.5 (1.39)	2.5 (1.41)	0.0 (0.99)	0.671	0.03
KIMRISS						
	Osteitis	34.4 (68.35)	53.7 (77.21)	19.3 (33.56)	0.066	0.57
	Synovitis/effusion	36.5 (16.1)	34.7 (12.59)	−1.8 (15.90)	0.756	0.11
Correlation MRI-WIPE vs. KIMRISS						
	Osteitis	0.75** (0.008)	0.97*** (<0.001)	0.85** (0.001)	–	–
	Synovitis/effusion	0.92*** (<0.001)	0.94*** (<0.001)	0.88*** (<0.001)	–	–
Exercise 4						
MRI-WIPE knee						
	Osteitis	1.5 (2.14)	1.7 (2.14)	0.2 (2.20)	0.720	0.10
	Synovitis/effusion	2.8 (1.50)	1.5 (0.91)	−1.4 (1.84)*	0.035	0.74
KIMRISS						
	Osteitis	11.0 (12.17)	16.8 (35.81)	5.8 (38.31)	0.463	0.15
	Synovitis/effusion	55.9 (22.06)	35.2 (13.28)	−20.7 (26.63)*	0.028	0.78
Correlation MRI-WIPE vs. KIMRISS						
	Osteitis	0.92*** (<0.001)	0.98*** (<0.001)	0.34 (0.332)	–	–
	Synovitis/effusion	0.89** (0.001)	0.67* (0.036)	0.89*** (<0.001)	–	–

Data are shown as mean (SD) and correlation coefficient (p-value). Comparisons of status scores at baseline and follow-up for cases with two timepoints were calculated with Wilcoxon signed-rank test. Spearman Rank Correlation analysis of MRI-WIPE variables versus KIMRISS variables were done for baseline and change. Standardized response mean (SRM) was calculated as mean change score divided by standard deviation (SD) of the change score and interpreted as follows: no: <0.20; small: ≥0.20 and <0.50; moderate: ≥0.50 and <0.80; large ≥0.80 [19]. $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$. KIMRISS, Knee Inflammation MRI Scoring System; MRI-WIPE, OMERACT MRI Whole-body score for Inflammation in Peripheral joints and Entheses in inflammatory arthritis; SRM, standardized response mean.

[1] Values are shown for the 3 readers with overall highest interreader agreement in exercise 3 (WPM, RGWL, MØ).

UCB, MSD and Pfizer; speaker/consultant for Pfizer, MSD, Novartis, BMS, AbbVie, UCB, Eli Lilly, Gilead and Celgene. HMO has received research grants from Janssen and Novartis; honoraria/speaker fees from AbbVie, Celgene, Janssen, Lilly, Novartis, Pfizer, Takeda and UCB. PGC has received speaker or consultancy fees from AbbVie, AstraZeneca, BMS, Eli Lilly, EMD Serono, Flexion Therapeutics, Galapagos, Gilead, Novartis, Pfizer and Stryker. MØ has received research support, consultancy fees and/or speaker fees from Abbvie, BMS, Boehringer-Ingelheim, Celgene, Eli-Lilly, Hospira, Janssen, Merck, Novartis, Pfizer, Regeneron, Roche, Sandoz, Sanofi and UCB. MW, IE, MSS, SK, VF, AJM, FG, JP, GDM, AEFP and JLJ have no declarations of interest for this work.

Acknowledgments

We thank CARE Arthritis Limited (carearthritis.com) for help with setting up the web-based scoring interface, the scoring exercises, and the web-based meetings. We thank all who participated in the SIG (Special Interest Group) virtual OMERACT meeting 29 October 2020. HMO, GDM and PGC are supported in part by the National Institute for Health Research (NIHR) Leeds Biomedical Research centre, United Kingdom. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix

Figs. A.1 and A.2

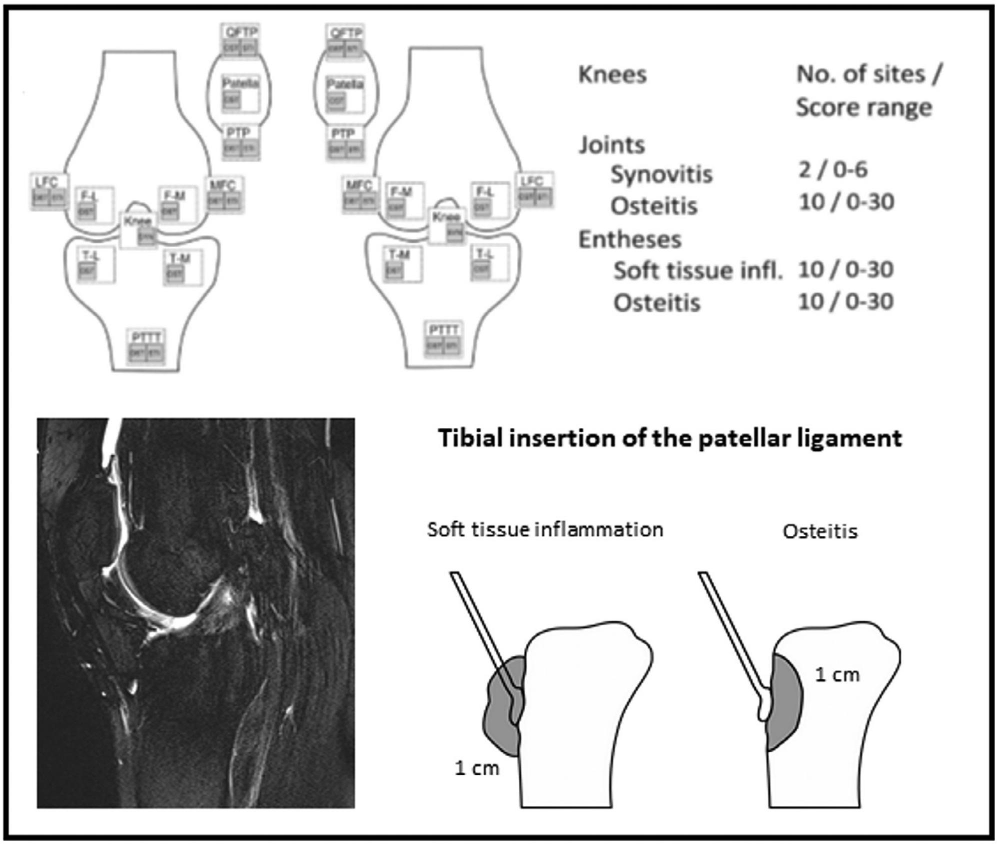

Fig. A.1. MRI-WIPE schematic and scoring ranges for the knee region (upper row), from Krabbe et al. [7]., and a schematic drawing of the principle of scoring osteitis and soft tissue inflammation using MRI-WIPE (lower row), illustrated with the tibial insertion of the patellar ligament (sagittal whole-body MR image of the knee region shown to the left). Using MRI-WIPE osteitis is assessed in the bone from the articular surface/entheseal insertion to a depth of 1 cm on all available images (as shown in the schematic of the tibial insertion of the patellar ligament). The osteitis 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: normal; 1: mild (1–33% of bone oedematous); 2: moderate (34–66% of bone oedematous); 3: severe (67–100% of bone oedematous). Soft tissue inflammation is assessed inside the ligament/tendon and its immediate surroundings to 1 cm from the entheseal insertion: 0: normal; 1: mild; 2: moderate; 3: severe – by thirds of the maximum potential volume of inflammatory tissue. Synovitis is assessed in the entire synovial compartment on all available images: 0: normal; 1: mild; 2: moderate; 3: severe – by thirds by thirds of the maximum potential volume of enhancing tissue in the synovial compartment [7].
F-L: femur-lateral; F-M: femur-medial; LFC: lateral femoral condyle; MFC: medial femoral condyle; MR: magnetic resonance; MRI-WIPE, OMERACT MRI Whole-body score for Inflammation in Peripheral joints and Entheses in inflammatory arthritis; OST, osteitis; PTP: patellar tendon insertion into patella; PTTT: patellar tendon insertion into tibial tuberosity; QFTP: quadriceps femoris tendon insertion into patella; STI, soft tissue inflammation; SYN, synovitis; T-L: tibia-lateral; T-M: tibia-medial.

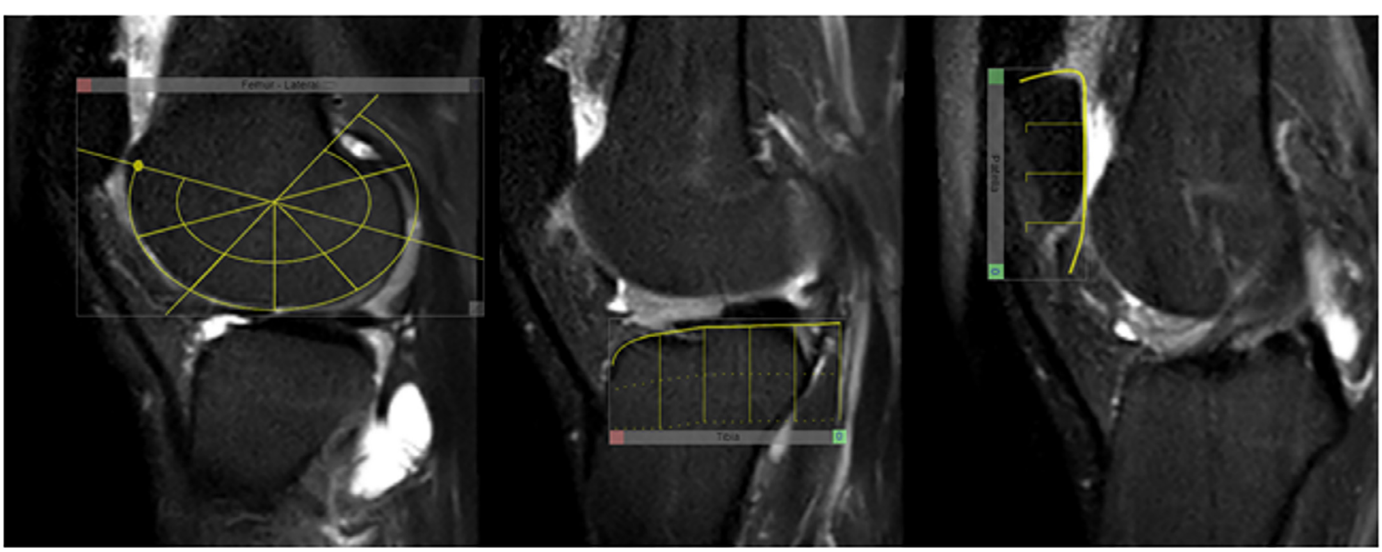

Fig. A.2. Sagittal whole-body MR image of a knee with the web-based superimposed interactive overlays positioned for femur, tibia and patella used in KIMRISS for osteitis scoring. Osteitis is scored on consecutive sagittal slices through the knee joint. The overlays are moved by the reader to fit bone at three sites for the femur and tibia (central slice and medial and lateral compartment). The position of the overlay is then automatically adjusted to best fit for other images slices. The overlay separates subarticular bone into approximately 1 × 1 cm regions. On each slice, the reader clicks each area with osteitis and sum scores of these regions are automatically calculated and adjusted for the scoring range of each region (total scoring range 0–500) [12, 13].
KIMRISS, Knee Inflammation MRI Scoring System; MR, magnetic resonance.

References

[1] Sudoł-Szopińska I, Matuszewska G, Kwiatkowska B, et al. Diagnostic imaging of psoriatic arthritis. Part I: etiopathogenesis, classifications and radiographic features. J Ultrason 2016;16(64):65–77.
[2] Sudoł-Szopińska I, Pracoń G. Diagnostic imaging of psoriatic arthritis. Part II: magnetic resonance imaging and ultrasonography. J Ultrason 2016;16 (65):163–74.
[3] Attard S, Castillo J, Zarb F. Establishment of image quality for MRI of the knee joint using a list of anatomical criteria. Radiography (Lond) 2018;24(3):196–203.
[4] Sudoł-Szopińska I, Jurik AG, Eshed I, et al. Recommendations of the ESSR arthritis subcommittee for the use of magnetic resonance imaging in musculoskeletal rheumatic diseases. Semin Musculoskelet Radiol 2015;19(4):396–411.
[5] Poggenborg RP, Eshed I, Ostergaard M, 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(5):823–9.
[6] Poggenborg RP, Pedersen SJ, Eshed I, 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 (Oxford) 2015;54(6):1039–49.
[7] Krabbe S, Eshed I, Gandjbakhch F, et al. Development and validation of an OMERACT MRI whole-body score for inflammation in peripheral joints and entheses in inflammatory arthritis (MRI-WIPE). J Rheumatol 2019;46(9):1215–21.
[8] Østergaard M, Eshed I, Althoff CE, 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 (11):1699–705.
[9] Boers M., Kirwan J.R., Tugwell P., Beaton D., Bingham C.O. III, Conaghan P.G., et al. The OMERACT Handbook 2018 [cited 2021 January 7]. Available from: https://www.omeract.org/handbook.
[10] Mathew AJ, Krabbe S, Eshed I, et al. The OMERACT MRI in enthesitis initiative: definitions of key pathologies, suggested MRI sequences, and a novel heel enthesitis scoring system. J Rheumatol 2019;46(9):1232–8.
[11] Ostergaard M, McQueen F, Wiell C, 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(8):1816–24.
[12] Jaremko JL, Jeffery D, Buller M, et al. Preliminary validation of the Knee Inflammation MRI Scoring System (KIMRISS) for grading bone marrow lesions in osteoarthritis of the knee: data from the Osteoarthritis Initiative. RMD Open 2017;3(1): e000355.
[13] [cited 2021 January 7]. Available from: https://www.carearthritis.com/mriportal/kimriss/.
[14] D'Agostino MA, Beaton DE, Maxwell LJ, Cembalo SM, Hoens AM, Hofstetter C, et al. Improving domain definition and outcome instrument selection: lessons learned for OMERACT from imaging. Semin Arthritis Rheum 2021 (In Press).
[15] Beaton DE, Maxwell LJ, Shea BJ, et al. Instrument selection using the OMERACT Filter 2.1: the OMERACT Methodology. J Rheumatol 2019;46(8):1028–35.
[16] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276–82.
[17] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15(2):155–63.
[18] Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23–34.
[19] Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53(5):459–68.