Validation of the OMERACT Magnetic Resonance Imaging Joint Space Narrowing Score for the Wrist in a Multireader Longitudinal Trial

Daniel Glinatsi, Siri Lillegraven, Espen A. Haavardsholm, Iris Eshed, Philip G. Conaghan, Charles Peterfy, Frédérique Gandjbakhch, Paul Bird, Pernille Bøyesen, Uffe M. Døhn, Harry K. Genant, and Mikkel Østergaard

ABSTRACT. ***Objective.*** To assess the intrareader and interreader agreement and sensitivity to change of the Outcome Measures in Rheumatology (OMERACT) Rheumatoid Arthritis Magnetic Resonance Imaging Joint Space Narrowing (RAMRIS-JSN) score in the rheumatoid arthritis (RA) wrist in a longitudinal multireader exercise.

Methods. Coronal T1-weighted MR image sets of 1 wrist from 20 patients with early RA were assessed twice for JSN at 17 sites at baseline and after 36 or 60 months by 4 readers blinded to patient data but not time order. The joints were scored 0–4 according to the OMERACT RAMRIS-JSN score. Intraclass correlation coefficients (ICC), smallest detectable change (SDC), percentage exact/close agreement (PEA/PCA), and standardized response mean (SRM) were calculated.

Results. Median baseline and change score was 10.3 and 1.9, respectively. Intrareader ICC for baseline and change scores was good (≥ 0.50) to very good (≥ 0.80) for all and 3 of 4 readers, respectively. Interreader ICC was very good for change (0.93), while poor for baseline score if all 4 readers were included (0.36), but very good if 1 reader was excluded (0.87). Intrareader and interreader SDC was low (2.34–3.18), except for the intrareader SDC for 1 reader (6.75). The mean PEA/PCA was high for baseline and change scores both within and between the readers (51.5–99.2), except for interreader baseline PEA (14.4). SRM was moderate for all readers (0.55–0.77).

Conclusion. The OMERACT RAMRIS-JSN score showed high overall intrareader and interreader reliability, and moderate sensitivity to change, supporting inclusion of the measure as part of the OMERACT RAMRIS system. (J Rheumatol First Release Feb 15 2015; doi:10.3899 /jrheum.141009)

Key Indexing Terms:

OMERACT RAMRIS JOINT SPACE NARROWING
RHEUMATOID ARTHRITIS MAGNETIC RESONANCE IMAGING

Because extended placebo-comparator studies are no longer ethical, distinguishing the difference between therapies in clinical trials has become more difficult[1]. Methods with high sensitivity are needed to detect change in patients with rheumatoid arthritis (RA).

The Outcome Measures in Rheumatology (OMERACT)

From the Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Glostrup Hospital, University of Copenhagen and Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; Department of Rheumatology, Diakonhjemmet Hospital, University of Oslo, Oslo, Norway; Sheba Medical Center, Tel Aviv University, Tel Aviv, Israel; Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and UK National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit, UK; Spire Sciences Inc., Boca Raton, Florida, USA; Department of Rheumatology, Pitié Salpêtrière Hospital, APHP, Université Paris 6-UPMC, Paris, France; University of New South Wales, Sydney, Australia; Medicine and Orthopedics, University of California, San Francisco, and Synarc Inc., San Francisco, California, USA.

D. Glinatsi, MD, Research Fellow, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Glostrup Hospital, University of Copenhagen; S. Lillegraven, MD, PhD, Postdoctoral Researcher, Department of Rheumatology, Diakonhjemmet Hospital, University of Oslo; E.A. Haavardsholm, MD, PhD, Postdoctoral Researcher, Department of Rheumatology, Diakonhjemmet Hospital; I. Eshed, MD, Professor of Radiology, Sheba Medical Center; P.G. Conaghan, MB, BS, PhD, FRACP, FRCP, Professor of Musculoskeletal Medicine, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, and NIHR Leeds Musculoskeletal Biomedical Research Unit; C. Peterfy, MD, PhD, FRCP, Chief Executive Officer, Spire Sciences Inc.; F. Gandjbakhch, MD, Practicing Rheumatologist, Department of Rheumatology, Pitié Salpêtrière Hospital, APHP, Université Paris 6-UPMC; P. Bird, BMed (Hons), FRACP, PhD, Grad Dip MRI, Senior Lecturer, University of New South Wales; P. Bøyesen, MD, PhD, Department of Rheumatology, Diakonhjemmet Hospital; U.M. Døhn, MD, PhD, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Glostrup Hospital, University of Copenhagen; H.K. Genant, MD, FACR, FRCR, Professor Emeritus of Radiology, Medicine and Orthopedics, University of California, San Francisco, and Synarc Inc.; M. Østergaard, MD, PhD, DMSc, Professor of Rheumatology, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Glostrup Hospital and Department of Clinical Medicine, University of Copenhagen.

Daniel Glinatsi and Siri Lillegraven contributed equally to this work.

Address correspondence to Dr. D. Glinatsi, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Glostrup Hospital, Entrance 5, st, DK-2600 Glostrup, Denmark. E-mail: daniel.glinatsi@gmail.com

RA Magnetic Resonance Imaging (MRI) Score (RAMRIS) assesses inflammation (synovitis and bone marrow edema) and bone damage (bone erosion), but when RAMRIS was introduced, cartilage damage was excluded because of insufficient image quality[2,3,4]. Nevertheless, MRI sequence technology has improved, making visualization of cartilage and JSN feasible.

Studies have shown that conventional radiography assessment of JSN in addition to bone erosion can provide valuable information on disease progression and impairment of physical function and work ability[5,6]. Further, the suppression of bone erosion by a drug may not correspond to the suppression of JSN[7]. Hence, adding MRI assessment of JSN may broaden the spectrum of clinically relevant structural damage pathologies evaluated in RA.

In 2011, the OMERACT MRI in arthritis working group presented a semiquantitative JSN scoring system for the hand as a potential addendum to RAMRIS (RAMRIS-JSN). Primary results have shown high intrareader and interreader reliability and moderate to good correlation with conventional radiography. Further good correlation was seen between MRI and computed tomography[8,9]. In the present multireader longitudinal study, we aimed at assessing the intrareader and interreader agreement and the sensitivity to change of the OMERACT RAMRIS-JSN score.

MATERIALS AND METHODS

Patients. Twenty RA patients from the Oslo Early RA cohort (disease duration < 1 year), all fulfilling the 1987 American College of Rheumatology classification criteria, were included in the exercise. All patients gave their informed consent.

Image acquisition. MRI of the dominant wrist was performed on a 1.5 Tesla MR scanner (General Electric, Signa) using a dedicated high-resolution wrist phased array coil. Coronal T1-weighted precontrast fast spin echo MR images (2.5 mm slice thickness, field of view 100 × 100 mm, matrix 320 × 256) were acquired. To obtain a wide range of JSN, MR images at baseline (all patients) and either at 36 months (2 patients) or 60 months' (18 patients) followup were selected by an assessor (SL) not participating in the reading exercise.

Scoring of images. The 20 paired image sets were read twice, on 2 consecutive days, on identical, dual-screen workstations (SECTRA PACS) by 4 experienced MRI readers familiar with the RAMRIS-JSN score (EH, IE, PC, MØ). A calibration session was performed using similar images the evening before the exercise. The readers were blinded for patient data but not for time order, as suggested by van Tuyl, *et al*[10]. The image sets were rerandomized and reanonymized before second reading. Images were scored at 17 sites in the wrist, according to the OMERACT RAMRIS-JSN score as described by Østergaard, *et al*[8]. Accordingly, JSN was defined as reduced joint space width compared to normal, as assessed in a slice perpendicular to the joint surface. Each site was given a score for narrowing between 0–4 as follows: 0, no narrowing; 1, focal or mild (< 33%); 2, moderate (34–66%); 3, moderate to severe (67–99%); 4, ankylosis.

Statistics. Status (baseline) scores and change (between baseline and followup) scores were calculated for each patient, in 2 independent readings by each reader. We used the mean of the 2 readings of the individual readers for calculations of interreader agreement and sensitivity to change.

Intrareader and interreader agreement was calculated using single measure and average measure intraclass correlation coefficients (ICC), respectively. An ICC ≥ 0.50 was considered good and an ICC ≥ 0.80 was considered very good. The smallest detectable change (SDC) was calculated for change scores[11] and was also expressed as the percentage of the highest score observed by the reader (mean maximal score between the readers for interreader SDC). Agreement between the readers was also expressed as the percentage of exact agreement (PEA), where the 4 readers had the same score, and as the percentage of close agreement (PCA), where the difference between the minimum and maximum score of the 4 readers was ≤ 1. Agreement between the first and second score for a single reader was also expressed as PEA and PCA.

Sensitivity to change was estimated using the standardized response mean (SRM: trivial, < 0.20; small, 0.20–0.49; moderate, 0.50–0.79; good, ≥ 0.80), calculated by dividing the mean change score with the SD of the change.

RESULTS

Baseline characteristics of the patients are presented in Table 1.

At baseline, JSN (total score ≥ 1) was found in all patients by all readers. The total JSN score at baseline of the individual readers was median 10.3 (mean 11.3), range 8.0–16.8, whereas the change in score from baseline to followup was median 1.9 (mean 2.1), range 1.5–3.0.

Intrareader and interreader ICC for total scores and SDC for total and individual joint space scores are presented in Table 2. Intrareader ICC was good to very good for all readers for baseline scores, and for 3 of 4 readers for change scores. Interreader ICC for baseline total scores was poor (0.36) when all 4 readers were included, but very good (0.87) when 1 reader was excluded. The interreader ICC was very good (0.93) for change scores. The intrareader SDC and the percentage of maximal score observed for total scores were low (i.e., good) for 3 of the 4 readers. Interreader SDC and percentage of mean maximal score observed were also low for total score. Intrareader and interreader SDC for separate joints were below 1 at all sites, i.e., below the increment of the scoring system.

The PEA and PCA for the individual joint spaces are presented in Table 3. Overall, the mean PEA on change score was 86.7% within readers and 51.5% between readers, while the corresponding mean PCA were 99.2% and 92.4%.

Table 1. Baseline characteristics of the 20 patients with rheumatoid arthritis included in the exercise.

Female, n (%)	15 (75)
Age, yrs, median (25th and 75th percentile)	55 (41;63)
Disease duration, days, median (25th and 75th percentile)	87 (36;136)
IgM RF positivity, n (%)	9 (45)
Anti-CCP positivity, n (%)	12 (60)
DAS28, median (25th and 75th percentile)	3.78 (3.30;4.72)
ESR, mm/h, median (25th and 75th percentile)	12 (10;21)
28-joint swollen joint count, median (25th and 75th percentile)	5.5 (4;10)

IgM RF: immunoglobulin M rheumatoid factor; CCP: cyclic citrullinated protein; DAS: Disease Activity Score; ESR: erythrocyte sedimentation rate.

Table 2. Intrareader and interreader intraclass correlation coefficients (ICC) and smallest detectable change (SDC).

	Intrareader				Interreader
	Reader 1	Reader 2	Reader 3	Reader 4	
ICC					
Baseline total score (95% CI)	0.89 (0.69 to 0.96)	0.79 (0.51 to 0.91)	0.52 (0.11 to 0.78)	0.96 (0.90 to 0.98)	0.36 (–0.03 to 0.68)
Total change score (95% CI)	0.41 (–0.02 to 0.72)	0.93 (0.83 to 0.97)	0.93 (0.83 to 0.97)	0.95 (0.87 to 0.98)	0.93 (0.87 to 0.97)
SDC (percentage of maximum score observed)					
Total score	6.75 (0.31)	2.56 (0.14)	2.34 (0.09)	3.18 (0.12)	2.88 (0.03)
CMC1	0.30 (0.15)	0.31 (0.15)	0.57 (0.38)	0.51 (0.34)	0.36 (0.05)
CMC2	0.31 (0.31)	0.31 (0.31)	0.68 (0.34)	0.55 (0.36)	0.29 (0.05)
CMC3	0.43 (0.21)	0.51 (0.34)	0.45 (0.30)	0.62 (0.41)	0.35 (0.05)
CMC4	0.62 (0.41)	0.68 (0.68)	0.93 (0.93)	0.57 (0.38)	0.26 (0.05)
CMC5	0.62 (1.24)	0.31 (0.31)	0.00 (NP*)	0.68 (0.68)	0.26 (0.11)
TRM-TRD	0.43 (0.85)	0.00 (NP*)	0.00 (NP*)	0.00 (NP*)	0.08 (0.15)
TRD-CAP	0.73 (0.73)	0.84 (0.56)	0.76 (0.51)	0.43 (0.43)	0.33 (0.07)
CAP-HAM	0.31 (0.15)	0.62 (0.62)	0.62 (0.25)	0.43 (0.28)	0.32 (0.05)
TRM-SCA	0.77 (0.51)	0.55 (0.27)	0.55 (0.22)	0.78 (0.31)	0.49 (0.06)
TRD-SCA	0.62 (0.62)	0.73 (0.48)	0.31 (0.21)	0.55 (0.55)	0.30 (0.06)
CAP-SCA	0.55 (0.27)	0.31 (0.21)	0.43 (0.21)	0.51 (0.17)	0.39 (0.05)
CAP-LUN	0.43 (0.21)	0.31 (0.15)	0.71 (0.35)	0.45 (0.22)	0.33 (0.04)
HAM-TRI	0.00 (0.00)	0.62 (0.41)	0.51 (0.20)	0.68 (0.27)	0.32 (0.04)
SCA-LUN	0.71 (0.28)	0.31 (0.15)	0.62 (0.21)	0.62 (0.25)	0.26 (0.03)
LUN-TRI	0.43 (0.21)	0.55 (0.27)	0.62 (0.21)	0.55 (0.22)	0.45 (0.05)
RAD-SCA	0.89 (0.89)	0.55 (0.55)	0.31 (0.31)	0.31 (0.15)	0.31 (0.06)
RAD-LUN	0.64 (0.32)	0.64 (0.25)	0.43 (0.17)	0.45 (0.15)	0.27 (0.03)

*NP: Not possible to calculate because of zero variance. Intrareader agreement is expressed as single measure ICC and interreader agreement as average measure ICC with a 95% CI. An ICC ≥ 0.50 is considered good and an ICC ≥ 0.80 is considered very good. SDC is calculated for intrareader and interreader change scores and is also expressed as the percentage of the maximal score observed by the reader. Interreader agreement is calculated by using the mean score of the 2 readings performed by each reader on a single timepoint. CMC: carpometacarpal; TRM: trapezium; TRD: trapezoid; CAP: capitate; HAM: hamate; SCA: scaphoid; LUN: lunate; TRI: triquetrum; RAD: radius.

Table 3. Percentage of exact agreement (PEA) and close agreement (PCA) for JSN assessment in each individual joint space.

	Intrareader Baseline		Intrareader Change		Interreader Baseline		Interreader Change	
	Mean PEA	Mean PCA	Mean PEA	Mean PCA	PEA	PCA	PEA	PCA
CMC1	73.8	100.0	87.5	100.0	5.0	75.0	45.0	100.0
CMC2	77.5	100.0	92.5	98.8	0.0	40.0	55.0	100.0
CMC3	66.3	95.0	88.8	100.0	0.0	60.0	50.0	85.0
CMC4	52.5	96.3	76.3	96.3	0.0	75.0	55.0	95.0
CMC5	61.3	100.0	93.8	100.0	5.0	70.0	55.0	100.0
TRM-TRD	92.5	97.5	97.5	100.0	65.0	100.0	90.0	100.0
TRD-CAP	58.8	91.3	70.0	98.8	15.0	50.0	35.0	95.0
CAP-HAM	60.8	98.7	87.3	98.7	0.0	50.0	60.0	95.0
TRM-SCA	73.8	100.0	81.3	100.0	15.0	60.0	30.0	70.0
TRD-SCA	70.0	95.0	84.8	100.0	0.0	45.0	35.0	90.0
CAP-SCA	78.8	100.0	90.0	100.0	10.0	85.0	55.0	85.0
CAP-LUN	80.0	100.0	91.3	97.5	5.0	80.0	60.0	95.0
HAM-TRI	79.7	100.0	87.3	100.0	20.0	85.0	60.0	90.0
SCA-LUN	82.5	100.0	90.0	97.5	45.0	95.0	45.0	100.0
LUN-TRI	73.8	100.0	83.8	100.0	15.0	75.0	40.0	75.0
RAD-SCA	77.5	100.0	87.5	98.8	15.0	85.0	45.0	100.0
RAD-LUN	82.5	100.0	85.0	100.0	30.0	90.0	60.0	95.0
All joints (mean)	74.1	96.0	86.7	99.2	14.4	71.8	51.5	92.4

Intrareader percentage of exact agreement (PEA) expresses the percentage of the patients receiving the same score twice by a single reader, and intrareader percentage of close agreement (PCA) is the percentage of the patients where the 2 readings differ no more than 1.0 in score. Intrareader PEA and PCA are presented as the mean of the 4 readers. Interreader PEA expresses the percentage of the patients receiving the same score by all 4 readers, and the interreader PCA is the percentage of the patients where the difference between the minimum and maximum score is ≤ 1.0. CMC: carpometacarpal; TRM: trapezium; TRD: trapezoid; CAP: capitate; HAM: hamate; SCA: scaphoid; LUN: lunate; TRI: triquetrum; RAD: radius.

Figure 1 displays the presence of baseline JSN (Figure 1A), change in JSN (Figure 1B), and mean JSN scores (Figure 1C and 1D), per individual joint space. The trapezium-scaphoid joint, trapezoid-capitate joint, and first carpometacarpal (CMC) joint showed change most frequently, while the CMC 5 and trapezium-trapezoid (TRM-TRD) joint showed the least change.

The SRM for total change scores was moderate in all readers (reader 1: 0.77; reader 2: 0.76; reader 3: 0.55; reader 4: 0.65).

DISCUSSION

In this longitudinal multireader early RA exercise the OMERACT RAMRIS-JSN score allowed assessment of JSN with overall high intrareader and interreader agreement for status and change scores. Further, a moderate SRM was found, indicating that the score allows detection of change in JSN.

The intrareader ICC for baseline and change scores was good to very good in all readers and in 3 readers, respectively, suggesting that the scoring system is reproducible. The low baseline interreader ICC in contrast to the very good interreader ICC for change score indicates that the readers were able to find similar change despite differences between the readers for status scores. It should be emphasized that high agreement for change is particularly relevant to clinical trials, because the change score will be used to assess potential differences in the effect of different interventions. It would be important to note that the interreader agreement for baseline total scores was very high if 1 reader was excluded. Future methodological studies similar to the present report may benefit from longer calibration sessions, including more image samples, to achieve better consistency between all readers. Previous studies of the OMERACT RAMRIS-JSN score have shown high intrareader and interreader ICC for status scores[9,10]. However, this was the first time, to our knowledge, that intrareader and interreader agreement were assessed longitudinally.

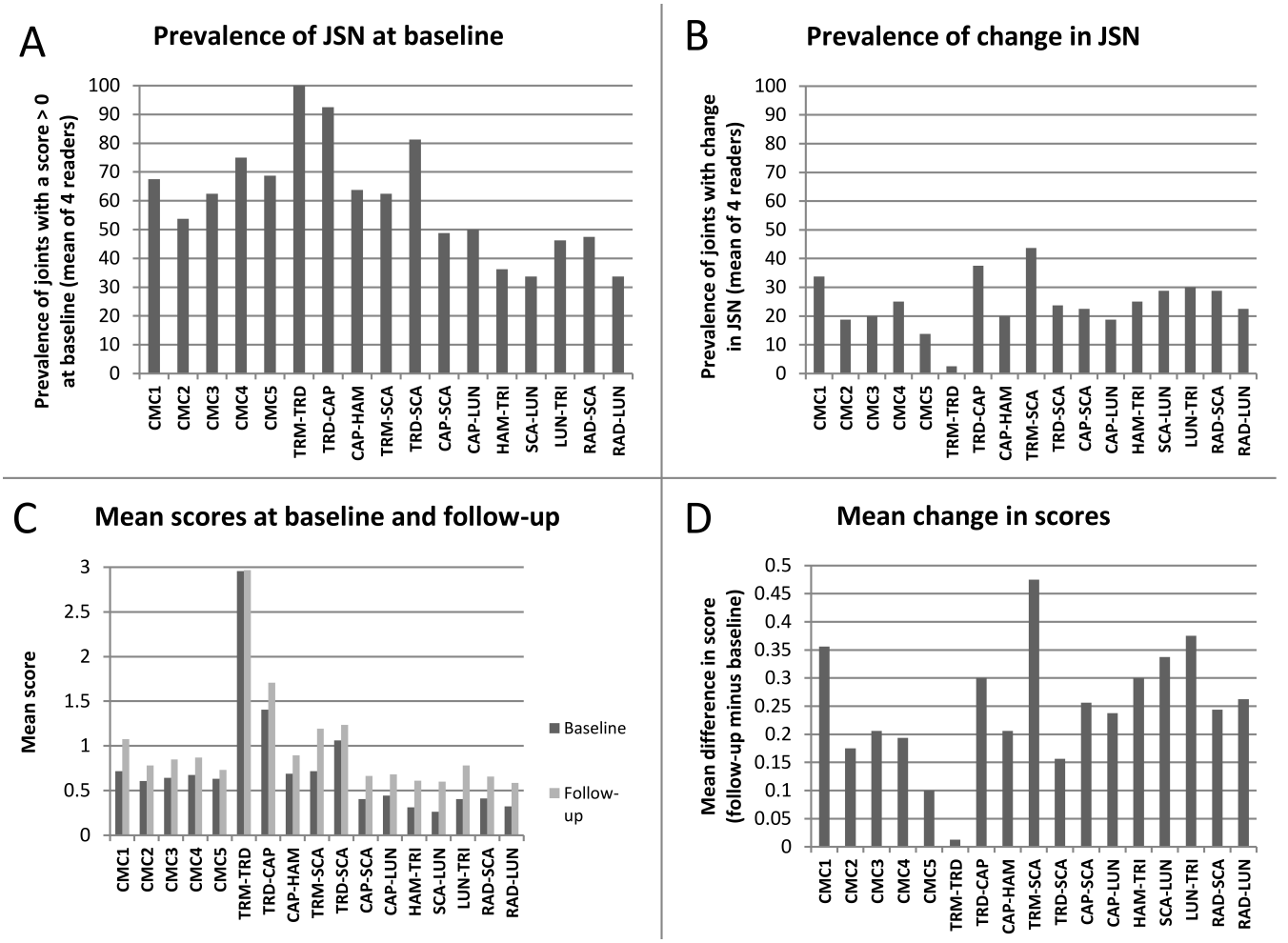

Figure 1. Prevalence and mean scores of JSN (status and change), per individual joint space. The prevalence of patients with a score > 0 at baseline (A) and for change scores (B) is shown at each individual joint site. Mean scores at baseline and followup (C) and for change scores (D) are means of the 8 reads (2 reads by each of the 4 readers). JSN: joint space narrowing; CMC: carpometacarpal; TRM: trapezium; TRD: trapezoid; CAP: capitate; HAM: hamate; SCA: scaphoid; LUN: lunate; TRI: triquetrum; RAD: radius.

The intrareader SDC for 3 readers and interreader SDC for all readers were low, suggesting that the OMERACT RAMRIS-JSN score is able to detect change in JSN. This is supported by the fact that the SDC represents only a low percentage of the maximal score observed. In addition, the SDC for separate joints were below 1 at all sites. This suggests that the OMERACT RAMRIS-JSN is able to reliably detect a change in the individual joint spaces at the level of the lowest increment of the scoring system.

Overall PEA and particularly PCA were high in the majority of joints for baseline scores, and high in all joints for change scores, despite the scoring exercise including 4 readers, making it difficult to reach high PEA. Higher agreement for change score may suggest that consistency is easier to achieve when assessing paired images than when assessing different cases, owing to anatomical variations. Some joint space sites, particularly the TRM-TRD joint, showed minimal change, which may support excluding this joint from the scoring system, despite high intrareader and interobserver agreement and high prevalence of baseline score above 0.

In the present study, the SRM for the total score was moderate in all readers, suggesting that the OMERACT RAMRIS-JSN score is sensitive to change. However, further evaluation of sensitivity to change in other populations, including unselected image sets and randomized controlled trials, is required.

The time between baseline and followup was 36–60 months in the present study. Future studies should investigate the change in JSN in modern clinical trials with short duration, because the change in JSN may be markedly less. However, patients selected for RA clinical trials often have high disease activity and more rapid development of joint damage than would be expected in a broad early RA cohort with moderate disease activity such as the present.

Other semiquantitative MRI scoring systems evaluating JSN of the RA hand have been reported. McQueen, *et al*[12] presented a 5-point scoring system assessing 8 joint spaces of the wrist. The scoring system showed high intrareader and interreader reliability for status scores, and 1 of 2 readers was able to detect a statistically significant change in JSN score over time[13]. However, not all joint spaces of the wrist were assessed; consequently it could not be assessed whether the selected areas were the most reliable and sensitive to change. Peterfy, *et al*[14,15] proposed a scoring system that evaluated articular cartilage directly rather than JSN. The method included the wrist joints, metacarpophalangeal (MCP), and proximal interphalangeal (PIP) joints using a 9-point scale, demonstrating high agreement with conventional radiography and a statistically significant change over time.

An advantage of the OMERACT RAMRIS-JSN score is that it can be incorporated as a part of the highly validated RAMRIS score, which assesses synovitis, bone marrow edema, and bone erosions; or it can be used separately. This may provide additional information on structural joint damage, which may be of importance to the physical function and work ability[5,6]. Further, the score of the JSN could easily be included in the RAMRIS evaluation, as no additional sequences would be required beyond those needed for erosion assessment. This allows assessment of JSN on preexisting MRI sets, in which coronal RAMRIS sequences have been obtained.

A limitation of the present study is the small sample size, and that the preselected cases do not represent a random selection, i.e., may not be representative of the general population. Further, chemical shift artefacts may complicate assessment of JSN when images without fat suppression are used, as in the present study. Finally, not blinding the readers to time order of the images may have introduced bias in favor of detecting progression of JSN, potentially influencing SRM calculations. A future step is to evaluate the sensitivity to change of the JSN score in a randomized controlled clinical trial.

The OMERACT RAMRIS-JSN score overall showed high intrareader and interreader reliability, and moderate sensitivity to change, supporting its validity as a tool for assessing JSN in RA clinical trials. This also supports that the measure can be included as a component of the OMERACT RAMRIS scoring system.

REFERENCES

1. O'Dell JR, Mikulus TR, Colbert RA, Weinblatt ME. American College of Rheumatology clinical trial priorities and design conference, July 22-23, 2010. Arthritis Rheum 2011;63:2151-6.
2. Ejbjerg B, McQueen F, Lassere M, Haavardsholm E, Conaghan P, O'Connor P, et al. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the wrist joint. Ann Rheum Dis Suppl 2005;64 Suppl 1:i23-47.
3. Ostergaard M, Klarlund M, Lassere M, Conaghan P, Peterfy C, McQueen F, et al. Interreader agreement in the assessment of magnetic resonance images of rheumatoid arthritis wrist and finger joints—an international multicenter study. J Rheumatol 2001;28:1143-50.
4. Ostergaard M, Peterfy C, Conaghan P, McQueen F, Bird P, Ejbjerg B, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. J Rheumatol 2003;30:1385-6.
5. Smolen JS, van der Heijde DM, Aletaha D, Xu S, Han J, Baker D, et al. Progression of radiographic joint damage in rheumatoid arthritis: independence of erosions and joint space narrowing. Ann Rheum Dis 2009;68:1535-40.
6. Smolen JS, van der Heijde DM, Keystone EC, van Vollenhoven RF, Goldring MB, Guerette B, et al. Association of joint space narrowing with impairment of physical function and work ability in patients with early rheumatoid arthritis: protection beyond disease control by adalimumab plus methotrexate. Ann Rheum Dis 2013;72:1156-62.
7. Cohen SB, Dore RK, Lane NE, Ory PA, Peterfy CG, Sharp JT, et al. Denosumab treatment effects on structural damage, bone mineral density, and bone turnover in rheumatoid arthritis: a twelve-month, multicenter, randomized, double-blind,

placebo-controlled, phase II clinical trial. Arthritis Rheum 2008;58:1299-309.
8. Ostergaard M, Boyesen P, Eshed I, Gandjbakhch F, Lillegraven S, Bird P, et al. Development and preliminary validation of a magnetic resonance imaging joint space narrowing score for use in rheumatoid arthritis: potential adjunct to the OMERACT RA MRI scoring system. J Rheumatol 2011;38:2045-50.
9. Dohn UM, Conaghan PG, Eshed I, Boonen A, Boyesen P, Peterfy CG, et al. The OMERACT-RAMRIS rheumatoid arthritis magnetic resonance imaging joint space narrowing score: intrareader and interreader reliability and agreement with computed tomography and conventional radiography. J Rheumatol 2014;41:392-7.
10. van Tuyl LH, van der Heijde D, Knol DL, Boers M. Chronological reading of radiographs in rheumatoid arthritis increases efficiency and does not lead to bias. Ann Rheum Dis 2014;73:391-5.
11. Bruynesteyn K, Boers M, Kostense P, van der Linden S, van der Heijde D. Deciding on progression of joint damage in paired films of individual patients: smallest detectable difference or change. Ann Rheum Dis 2005;64:179-82.
12. McQueen F, Clarke A, McHaffie A, Reeves Q, Williams M, Robinson E, et al. Assessment of cartilage loss at the wrist in rheumatoid arthritis using a new MRI scoring system. Ann Rheum Dis 2010;69:1971-5.
13. McQueen FM, McHaffie A, Clarke A, Lee AC, Reeves Q, Curteis B, et al. MRI osteitis predicts cartilage damage at the wrist in RA: a three year prospective 3T-MRI study examining cartilage damage. Arthritis Res Ther 2014;16:R33.
14. Peterfy CG, DiCarlo JC, Olech E, Bagnard MA, Gabriele A, Gaylis N. Evaluating joint-space narrowing and cartilage loss in rheumatoid arthritis by using MRI. Arthritis Res Ther 2012;14:R131.
15. Peterfy CG, Olech E, Dicarlo JC, Merrill JT, Countryman PJ, Gaylis NB. Monitoring cartilage loss in the hands and wrists in rheumatoid arthritis with magnetic resonance imaging in a multi-center clinical trial: IMPRESS (NCT00425932). Arthritis Res Ther 2013;15:R44.