The Longitudinal Reliability and Responsiveness of the OMERACT Hand Osteoarthritis Magnetic Resonance Imaging Scoring System (HOAMRIS)

Ida K. Haugen, Iris Eshed, Frederique Gandjbakhch, Violaine Foltz, Mikkel Østergaard, Pernille Bøyesen, Paul Bird, Harry K. Genant, Charles G. Peterfy, and Philip G. Conaghan

ABSTRACT. ***Objective.*** To evaluate the interreader reliability of change scores and the responsiveness of the OMERACT Hand Osteoarthritis (OA) Magnetic Resonance Image (MRI) Scoring System (HOAMRIS).

Methods. Paired MRI (baseline and 5-yr followup) from 20 patients with hand OA were scored with known time sequence by 3 readers according to the HOAMRIS: Synovitis, erosive damage, cysts, osteophytes, cartilage space loss, malalignment, and bone marrow lesions (BML; 0–3 scales with 0.5 increments for synovitis, erosive damage, and BML). Interreader reliability for status and change scores were assessed by intraclass correlation coefficients (ICC), percentage exact agreement and percentage close agreement (PEA/PCA), and smallest detectable change (SDC). Responsiveness was assessed by standardized response means (SRM).

Results. Cross-sectional interreader ICC were good to very good (≥ 0.74) for all features except synovitis, cysts, and malalignment (ICC 0.50–0.58). The range of change values was small, leading to low ICC for change scores. The SDC values for sum scores (total range 0–24) varied between 1.97–3.05 (except 1.08 for malalignment). For status scores, PEA/PCA on scores in individual joints across the readers were 8.1–50.0 and 43.8–78.1, respectively. Similarly, PEA/PCA for change scores were 20.6-63.8 and 66.3-93.1, respectively. All features except cysts and BML demonstrated good responsiveness with higher SRM for sum scores (range 0.46–1.62) than for scores in individual joints (range 0.24–0.73).

Conclusion. Good to very good interreader ICC values were found for cross-sectional readings, whereas the longitudinal reliability was lower because of a smaller range of change scores. All features, except cysts and BML, showed good responsiveness. (First Release September 1 2015; J Rheumatol 2015;42:2486–91; doi:10.3899/jrheum.140983)

Key Indexing Terms:
OMERACT HAND OSTEOARTHRITIS MAGNETIC RESONANCE IMAGING OUTCOMES

From the Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway; Department of Diagnostic Imaging, Sheba Medical Center, Tel Aviv University, Tel Aviv, Israel; Department of Rheumatology, Pitié Salpêtriere Hospital, APHP, Université Pierre et Marie Curie, Paris, France; Copenhagen Center for Arthritis Research, Center for Rheumatology and Spinal Diseases, Glostrup Hospital, Glostrup, Denmark; Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark; Department of Rheumatology, St. George Hospital, Sydney, Australia; St. George Clinical School, University of New South Wales, Sydney, Australia; Departments of Radiology and Medicine, University of California San Francisco, San Francisco, California, USA; Synarc Inc., Newark, California, USA; Spire Sciences Inc., Boca Raton, Florida, USA; Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and UK National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK.

I.K. Haugen has received funding from the Norwegian Rheumatology Foundation (Norsk Revmatikerforbund)/Ekstrastiftelsen. Data collection in the Oslo hand OA cohort is supported by grants from Grethe Harbitz' Legacy For Combating Rheumatic Diseases and the Dr. Trygve Gythfeldt og Wife's Research Fund.

I.K. Haugen, MD, PhD, Postdoctoral Researcher, Department of Rheumatology, Diakonhjemmet Hospital; I. Eshed, MD, Professor of Radiology, Sheba Medical Center, Tel Aviv University; F. Gandjbakhch, MD, Practicing Rheumatologist, Department of Rheumatology, Pitié Salpêtriere Hospital, APHP, Université Pierre et Marie Curie; V. Foltz, MD, Practicing Rheumatologist, Department of Rheumatology, Pitié Salpêtriere Hospital, APHP, Université Pierre et Marie Curie; M. Østergaard, MD, PhD, DMSc, Professor of Rheumatology, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spinal Diseases, Glostrup Hospital, and Department of Clinical Medicine, University of Copenhagen; P. Bøyesen, MD, PhD, Postdoctoral Researcher, Department of Rheumatology, Diakonhjemmet Hospital; P. Bird, BMed (Hons), FRACP, PhD, Practicing Rheumatologist, Department of Rheumatology, St. George Hospital and Senior Lecturer, University of New South Wales; H.K. Genant, MD, FACR, FRCR, Professor Emeritus of Radiology, Medicine and Orthopedics, University of California San Francisco, and Senior Consultant, Synarc Inc.; C.G. Peterfy, MD, PhD, FRCPC, President and Chief Executive Officer, Spire Sciences Inc.; 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.

Address correspondence to Dr. I.K. Haugen, Department of Rheumatology, Diakonhjemmet Hospital, Post Box 23 Vinderen, 0319 Oslo, Norway. E-mail: haugen_ida@hotmail.com

Osteoarthritis (OA) is a "whole-joint" disease[1], and magnetic resonance imaging (MRI) has the ability to visualize all affected joint structures[2]. Previous hand OA studies have shown that MRI is more sensitive than radiographs in detection of structural features[3,4]. Further, MRI findings such as synovitis and bone marrow lesions (BML) are associated with joint tenderness[5]. Hence, MRI is a valuable tool to increase the understanding of the pathogenesis of OA, and in future clinical trials may serve as an important outcome measure.

The Oslo hand OA MRI scoring system included assessment of several structural (osteophytes, joint space narrowing, erosions, cysts, BML, malalignment, collateral ligament pathology) and inflammatory (synovitis and flexor tenosynovitis) features in the distal interphalangeal (DIP) and proximal interphalangeal (PIP) joints[6]. Despite good to very good reliability[4,5,6], there were several limitations. Because of inclusion of many features, and because scoring of the proximal and distal parts of the joint was done separately, the scoring system was time-consuming. Further, features such as collateral ligament pathology and flexor tenosynovitis were uncommon, had lower reliability compared with other features, and no associations were found with joint tenderness[4,5,6].

Based on this tool, the OMERACT MRI Working Group iteratively developed a preliminary OMERACT hand OA MRI scoring system (OMERACT HOAMRIS) using OMERACT methodology[7]. In a reliability exercise with cross-sectional readings, the interreader reliability was good to very good for all features[7].

To develop a satisfactory tool for clinical trials, assessment of the reliability of change scores and the responsiveness of HOAMRIS is warranted. However, longitudinal hand OA MRI studies are uncommon. We assessed the reliability and the responsiveness of status and change scores of the HOAMRIS in 20 patients from the Oslo hand OA cohort with hand MRI scans at 2 timepoints 5 years apart.

MATERIALS AND METHODS

Three readers including 1 radiologist (IE) and 2 rheumatologists (FG, VF) participated in the reliability exercise. All readers participated in the previous hand OA MRI exercise[7].

Patients in the Oslo hand OA cohort had MRI scans of the DIP and PIP joints of the dominant hand acquired with a 1.0T extremity scanner (ONI, GE Healthcare) at 2 examinations (2008–2009 and 2013). MRI sequences included short-tau inversion recovery images in coronal and axial planes (TE 16.3 and 21 ms, TR 2850 and 3150 ms, slice thickness 2–3 mm, gap between slices 0.2 and 1 mm) and T1-weighted gradient-echo fat-suppressed pre- and post-gadolinium images in coronal, axial, and sagittal planes (TE 5 ms, TR 20 ms, slice thickness 1 mm, gap between slices 0 mm).

Data collection in the Oslo hand OA cohort was approved by the regional ethics committee and the data inspectorate. All patients signed informed consent.

Calibration Exercise

A calibration exercise was performed in March 2014. Three readers (IE, FG, VF) each scored 3 patients from the Oslo hand OA cohort (1 timepoint) according to the proposed HOAMRIS. The HOAMRIS includes assessment of synovitis, erosive damage, cysts, osteophytes, cartilage space loss, malalignment, and BML (all features on 0–3 scales with 0.5 increments for synovitis, erosions and BML; see Appendix 1, which was presented in the previous publication)[7]. After reading, scoring discrepancies were evaluated by a Web-based meeting.

An updated and extended version of the atlas was distributed and approved by all readers prior to the reliability exercise. The coronal plane was recommended for evaluation of all MRI features, except synovitis, for which we agreed to use the axial plane for standardization purposes (as previously proposed). Both coronal and sagittal planes were used for assessment of osteophytes. The scoring system (including definitions, grading, and recommended planes) was not changed as compared to the original publication[7].

Reliability and Responsiveness Exercise

An interreader reliability exercise was performed between April and May 2014. Each reader scored 20 patients who had MRI scans at 2 timepoints. Paired MRI scans were read with known time sequence using the HOAMRIS. MRI scans were selected by a nonreader based on availability of appropriate sequences, wide range of progression of radiographic hand OA structural severity [based on changes in Kellgren-Lawrence (KL) scores] and a wide range of changes in clinical inflammatory features (based on changes in swollen joint counts).

Statistical Analysis

We calculated the median and interquartile range for each MRI feature based on the reader mean values.

Interreader reliability was calculated for status and change scores. Percentage exact agreement (PEA), percentage close agreement (PCA) and average measure intraclass correlation coefficients (ICC) were calculated using mixed-effect models (absolute agreement). To determine whether the change of sum score in an individual patient was beyond the measurement error, we calculated the smallest detectable change (SDC), which is 1.96 times the standard error of measurement of the change score (based on the residual error) divided by the square root of the number of readers (in this case, 3)[8]. PEA was defined as a difference of 0 or 0.5 between the minimum and maximum scores across the 3 readers, whereas PCA was defined as a difference of ≤ 1 between the minimum and maximum scores. ICC values < 0.20 were considered as poor reliability, 0.20 ≤ ICC < 0.40 as fair, 0.40 ≤ ICC < 0.60 as moderate, 0.60 ≤ ICC < 0.80 as good, and 0.80 ≤ ICC < 1.00 as very good reliability[9].

Responsiveness was assessed by standardized response means (SRM) at joint level and patient level (i.e., sum scores for the 8 DIP and PIP joints). The change score for each patient was averaged across the 3 readers ("averaged change score"). Thereafter, the SRM were computed by dividing the mean "averaged change scores" by the SD of the "averaged change scores." SRM values ≥ 0.80 were considered as good responsiveness, 0.50 ≤ SRM < 0.80 as moderate, and SRM < 0.50 as low responsiveness[10].

RESULTS

The demographic and clinical variables for the 20 patients are presented in Table 1. Among the 160 assessed joints, radiographic hand OA was present in 119 [74.4%; KL grade 2 in 58 (36.3%) joints, KL grade 3 in 31 (19.4%) joints, and KL grade 4 in 30 (18.8%) joints]. The mean (SD) followup time between the 2 MRI examinations was 4.6 (0.3) years. During followup, we observed progression of all features (range 2.0–3.0 for sum scores for most features except malalignment and BML; Table 2). For synovitis, BML, and cysts, decrease during followup was observed by 1 or more reader(s) in 46/160 (28.8%) joints, whereas increasing synovitis, cysts, and BML scores were observed by 1 or more

Table 1. Baseline characteristics of the 20 patients with hand OA included in the exercise.

Baseline Characteristics	
Women, n (%)	19 (95)
Age, mean (SD), yrs	65.8 (4.5)
Body mass index, mean (SD), kg/m^2 *	26.5 (3.9)
ACR criteria for hand OA, n (%)	20 (100)
KL sum (DIP/PIP dominant hand), mean (SD) [range 0–32]	17.3 (5.0)
No. swollen joints (DIP/PIP dominant hand), mean (SD) [range 0–8]	4.0 (1.9)
AUSCAN pain, mean (SD) [0–20]	8.9 (3.1)
AUSCAN physical, mean (SD) [0–36]	18.1 (6.6)
Grip strength dominant hand, mean (SD) kg	20.1 (6.9)

* Body mass index data on 19 patients (one missing). ACR: American College of Rheumatology; KL: Kellgren-Lawrence; DIP: distal interphalangeal; PIP: proximal interphalangeal; AUSCAN: Australian/Canadian hand index; OA: osteoarthritis.

Table 2. The median status scores and change scores for the OMERACT hand OA MRI score.

	Median (IQR) at Baseline*	Median (IQR) at Followup*	Median (IQR) Change*
Joint level (0–3 scale)			
Synovitis	0.7 (0.3, 1.0)	1.0 (0.7, 1.3)	0.3 (0.0, 0.7)
Erosive damage	0.7 (0.0, 1.3)	1.0 (0.3, 2.0)	0.2 (0.0, 0.7)
Cysts	0.3 (0.0, 0.7)	0.7 (0.3, 1.3)	0.3 (0.0, 0.7)
Osteophytes	1.0 (0.7, 1.7)	1.3 (0.7, 2.0)	0.1 (0.0, 0.3)
Cartilage space loss	1.0 (0.3, 2.0)	1.3 (0.7, 2.3)	0.3 (0.0, 0.6)
Malalignment	0.3 (0.0, 0.7)	0.2 (0.0, 0.9)	0.0 (0.0, 0.3)
Bone marrow lesions	0.3 (0.0, 0.7)	0.3 (0.0, 1.0)	0.0 (0.0, 0.3)
Patient level (sum scores for 8 DIP and PIP joints, 0–24 scales)			
Synovitis	5.3 (4.1, 7.1)	7.0 (6.0, 10.0)	2.3 (1.1, 3.3)
Erosive damage	6.5 (2.1, 9.9)	9.2 (3.7, 14.4)	2.5 (0.8, 4.8)
Cysts	4.2 (2.2, 6.7)	6.3 (3.2, 9.0)	2.0 (0.0, 4.6)
Osteophytes	11.0 (4.8, 13.7)	13.0 (7.6, 16.7)	3.0 (1.0, 3.8)
Cartilage space loss	10.0 (3.8, 13.9)	11.3 (5.3, 17.5)	2.0 (1.3, 3.7)
Malalignment	2.8 (1.1, 4.7)	4.2 (2.1, 5.9)	0.8 (0.4, 1.3)
Bone marrow lesions	3.0 (1.2, 4.7)	4.5 (2.0, 6.3)	0.6 (–0.3, 2.0)

*Median and interquartile range (IQR) for each of the MRI features are based on average score for the 3 readers. OA: osteoarthritis; MRI: magnetic resonance imaging; DIP: distal interphalangeal; PIP: proximal interphalangeal.

reader(s) in 108/160 (67.5%), 114/160 (71.3%), and 86/160 (53.8%) joints, respectively.

Reliability

The cross-sectional interreader ICC values were good to very good for erosive damage, osteophytes, cartilage space loss, and BML, whereas reliability was moderate for synovitis, cysts, and malalignment. Close agreement was found in ≥ 43.8% of the joints for all MRI features (lowest value for osteophytes), whereas the exact agreement was generally low (Table 3).

Table 3. Cross-sectional interreader reliability of the OMERACT hand OA MRI score.

	AvmICC (95% CI)	PEA	PCA
Baseline (2008–09)			
Synovitis	0.50 (0.05, 0.78)	18.8	68.8
Erosive damage	0.81 (0.32, 0.93)	31.3	59.4
Cysts	0.51 (0.02, 0.78)	30.6	67.5
Osteophytes	0.74 (0.19, 0.91)	10.0	44.4
Cartilage space loss	0.93 (0.72, 0.98)	20.6	55.6
Malalignment	0.58 (0.10, 0.82)	50.0	76.9
Bone marrow lesions	0.87 (0.72, 0.94)	49.4	78.1
Followup (2013)			
Synovitis	0.57 (0.09, 0.82)	14.4	55.0
Erosive damage	0.84 (0.33, 0.95)	22.5	51.9
Cysts	0.53 (0.02, 0.80)	25.6	51.9
Osteophytes	0.76 (0.13, 0.92)	8.1	43.8
Cartilage space loss	0.92 (0.71, 0.97)	16.3	60.0
Malalignment	0.51 (0.00, 0.78)	40.0	69.4
Bone marrow lesions	0.91 (0.79, 0.97)	40.6	74.4

OA: osteoarthritis; MRI: magnetic resonance imaging; AvmICC: average measure intraclass correlation coefficients; PEA: percentage exact agreement (joint level); PCA: percentage close agreement (joint level).

The interreader ICC values for change scores were good for erosive damage and BML and moderate for cysts and osteophytes. Poor to fair reliability was observed for synovitis, cartilage space loss, and malalignment. For most features (except malalignment), the SDC values varied between 1.97 and 3.05 (Table 4). Close agreement was found in ≥ 66.3% of the joints for all MRI features (lowest for synovitis, BML, and cysts). The exact agreement ranged from 20.6% to 63.8% for synovitis and malalignment, respectively (Table 4).

In general, the PIP joints demonstrated higher exact agreement in the cross-sectional evaluation of synovitis, cysts, malalignment, and BML than the DIP joints (data not shown). For change scores, the exact agreement was higher for cysts, osteophytes, malalignment, and BML; and close

Table 4. The interreader reliability of change scores of the OMERACT hand OA MRI score.

	AvmICC (95% CI)	SDC	PEA	PCA
Synovitis	–0.07 (–0.89, 0.49)	3.05	20.6	66.3
Erosive damage	0.74 (0.47, 0.89)	2.08	44.4	76.3
Cysts	0.59 (0.18, 0.82)	2.98	28.8	67.5
Osteophytes	0.53 (0.08, 0.79)	1.97	46.9	80.0
Cartilage space loss	0.08 (–1.02, 0.62)	2.71	37.5	82.5
Malalignment	0.39 (–0.07, 0.71)	1.08	63.8	93.1
Bone marrow lesions	0.76 (0.51, 0.90)	2.56	38.8	66.9

OA: osteoarthritis; MRI: magnetic resonance imaging; AvmICC: average measure intraclass correlation coefficients; SDC: smallest detectable change; PEA: percentage exact agreement (joint level); PCA: percentage close agreement (joint level).

agreement was higher for synovitis in the PIP joint than in the DIP joints (data not shown).

Higher cross-sectional ICC values were found for all features except synovitis and BML in patients with mild to moderate radiographic hand OA at baseline (n = 9 with KL sum score 8–16) compared to patients with more severe radiographic hand OA (n = 11 with KL sum score 19–26; data not shown). The changes were larger for the majority of MRI features (except for similar degree of change for cartilage space loss) in the group with more severe radiographic hand OA. Higher ICC values for change scores were observed for synovitis, cartilage space loss, malalignment, and BML in patients with severe radiographic OA, whereas the reliability of changes in cysts was better in mild disease. There were no differences in ICC values for osteophytes and erosions (data not shown).

Responsiveness

The SRM values for the MRI sum scores were good for all features except cysts and BML, which showed moderate and low SRM values, respectively (Table 5). Analyzing the responsiveness for the MRI sum scores for each reader separately revealed lower SRM values (data not shown). The SRM values for MRI features at individual joint level were low to moderate (Table 5).

DISCUSSION

In our study, the OMERACT MRI working group tested the reliability of status scores and change scores and the responsiveness of the HOAMRIS.

As previously shown[7], good reliability was found for the cross-sectional readings. Compared to the previous exercise[7], the interreader ICC values were slightly lower despite the same readers. However, the patients in the current study had generally more severe disease when comparing the median values for the MRI features. Joints with severe structural abnormalities may be more difficult to score, which may decrease reliability. In the current study, we found higher cross-sectional reliability for all features except synovitis and BML in patients with mild to moderate radiographic hand OA, supporting our hypothesis.

Table 5. Responsiveness of the OMERACT hand OA MRI score.

	SRM at Joint Level	SRM at Patient Level (sum scores)
Synovitis	0.58	1.49
Erosive damage	0.62	1.24
Cysts	0.46	0.79
Osteophytes	0.63	1.57
Cartilage space loss	0.73	1.62
Malalignment	0.52	1.22
Bone marrow lesions	0.24	0.46

OA: osteoarthritis; MRI: magnetic resonance imaging; SRM: standardized response mean.

Whereas the PEA and PCA values were similar for both status and change scores, we found considerably lower interreader ICC values for change scores compared to the ICC values for status scores. This may be explained by a smaller range for change scores leading to lower total variance. Further, in joints without close agreement, the magnitude of the difference was larger for change scores than for status scores (i.e., both increase and decrease were possible for change scores). In contrast to cross-sectional reliability, which overall was higher in patients with milder disease, we found higher reliability for change scores of several MRI features in patients with more severe disease. These differences are most likely due to larger changes of pathology in patients with severe OA, whereas more subtle changes in patients with milder disease were difficult to assess reliably.

Low reliability may further be explained by the small size of the joints, insufficient pre-exercise training, and the quality of the 1.0T MRI scans. Using these scans, the distinction between marginal erosions versus cysts and cysts versus BML may be difficult, emphasizing the need for validation studies using CT and/or histology. In addition, scoring is complicated by the small size of the joints. In general, we found higher reliability for the PIP joints than the smaller DIP joints, as shown in the Oslo hand OA cohort[6]. Further, the readers used different software and screens of different sizes, which may also decrease reliability. The 20 patients with hand OA in this exercise were carefully selected based on the amount of radiographic OA progression and changes in swollen joints. Hence, there was a spectrum of pathology for all MRI features. Therefore the relatively low number of 20 patients is not the major cause of the poor reliability observed for certain features.

Responsiveness was evaluated by calculation of SRM values. In general, the SRM values were good for most features (except cysts and BML) using the sum score, as opposed to looking at responsiveness in individual joints. Hence, in clinical trials, sum scores may be the most responsive measure. It must be noted, however, that not blinding the readers to time order of the images could have introduced a bias in favor of progression. BML and cysts showed lower responsiveness, which is at least partly related to the natural history of these features. As shown in knee OA[11,12,13], BML and cysts were frequently decreased. Further, BML may develop into cysts[14]. Decrease of synovitis was also frequently present, as shown by Kortekaas, *et al*[15]. However, high SRM was observed for synovitis. Other factors affecting responsiveness may relate to reliability and image quality. The current study included mostly women, because the majority of participants in the Oslo hand OA cohort are women. Women may experience a larger progression of hand OA features than men[16], and therefore a better responsiveness.

In the Oslo hand OA cohort, only the DIP and PIP joints were covered by the coil, and imaging of the thumb base joint

would have required a separate acquisition. In future studies, the definitions of HOAMRIS will need to be validated for the thumb base, which is frequently affected by OA and is important for both pain and function[17].

Our results are in line with a previous MRI study by Haugen, *et al*, showing a high prevalence of MRI features including synovitis in patients with hand OA[4], emphasizing that hand OA is a disease of the whole joint with a substantial inflammatory component. With respect to future clinical studies, our results suggest that MRI may be a responsive tool in hand OA clinical trials. However, calibration of readers should be carefully done in all studies to optimize the reliability of readings.

The ICC values for change scores were especially low for synovitis and cartilage space loss. Both features represent important features of OA, and should probably be addressed in clinical OA trials. Using the current MRI sequences, we were not able to assess the cartilage directly in these small finger joints. Hence, cartilage space loss was used as a marker similar to joint space narrowing on conventional radiographs. A previous study has demonstrated higher sensitivity of conventional radiography in detection of joint space narrowing[4], and may be a better imaging modality to assess cartilage until we have MR images with more optimal resolution and sequences. Synovitis will probably be an important outcome in future clinical trials. Other assessment tools such as clinical examination and ultrasound are also hampered by modest reliability. MRI may represent the most promising tool to detect changes in inflammation owing to pairwise comparisons of images. Because of the modest reliability for change scores in the current study, we recommend more intensive training of the readers as well as better MRI quality to facilitate better interreader reliability. Future studies should confirm the high responsiveness and reevaluate the longitudinal reliability using MRI scans of higher quality. Currently, the OMERACT MRI working group does not have available data from any observational longitudinal cohorts or clinical trials using 1.5T or 3.0T MRI as an outcome measure.

Our results suggest that MRI is sensitive to change in hand OA. For cross-sectional readings the reliability was good, with high ICC values. The range of change scores was small, leading to lower ICC values for change scores. Further validation of MRI measurements is needed.

ACKNOWLEDGMENT

We thank Barbara Slatkowsky-Christensen for her contribution in data collection in the Oslo hand OA cohort, and Siri Lillegraven, Espen A. Haavardsholm, Tore K. Kvien, Sølve Sesseng, and Désirée van der Heijde for their important contributions in the development of the Oslo hand OA MRI scoring system.

REFERENCES

1. Abramson SB, Attur M. Developments in the scientific understanding of osteoarthritis. Arthritis Res Ther 2009;11:227.
2. Guermazi A, Roemer FW, Hayashi D. Imaging of osteoarthritis: update from a radiological perspective. Curr Opin Rheumatol 2011;23:484-91.
3. Wittoek R, Jans L, Lambrecht V, Carron P, Verstraete K, Verbruggen G. Reliability and construct validity of ultrasonography of soft tissue and destructive changes in erosive osteoarthritis of the interphalangeal finger joints: a comparison with MRI. Ann Rheum Dis 2011;70:278-83.
4. Haugen IK, Bøyesen P, Slatkowsky-Christensen B, Sesseng S, Bijsterbosch J, van der Heijde D, et al. Comparison of features by MRI and radiographs of the interphalangeal finger joints in patients with hand osteoarthritis. Ann Rheum Dis 2012;71:345-50.
5. Haugen IK, Bøyesen P, Slatkowsky-Christensen B, Sesseng S, Van Der Heijde D, Kvien TK. Associations between MRI-defined synovitis, bone marrow lesions and structural features and measures of pain and physical function in hand osteoarthritis. Ann Rheum Dis 2012;71:899-904.
6. Haugen IK, Lillegraven S, Slatkowsky-Christensen B, Haavardsholm EA, Sesseng S, Kvien TK, et al. Hand osteoarthritis and MRI: development and first validation step of the proposed Oslo Hand Osteoarthritis MRI score. Ann Rheum Dis 2011;70:1033-8.
7. Haugen IK, Ostergaard M, Eshed I, McQueen FM, Bird P, Gandjbakhch F, et al. Iterative development and reliability of the OMERACT Hand Osteoarthritis MRI Scoring System. J Rheumatol 2014;41:386-91.
8. 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.
9. Müller R, Büttner P. A critical discussion of intraclass correlation coefficients. Stat Med 1994;13:2465-76.
10. Cohen R. Statistical power analysis for the behavioral sciences. 2nd edition. London: Routledge; 2013.
11. Tanamas SK, Wluka AE, Pelletier J-P, Martel-Pelletier J, Abram F, Wang Y, et al. The association between subchondral bone cysts and tibial cartilage volume and risk of joint replacement in people with knee osteoarthritis: a longitudinal study. Arthritis Res Ther 2010;12:R58.
12. Xu L, Hayashi D, Roemer FW, Felson DT, Guermazi A. Magnetic resonance imaging of subchondral bone marrow lesions in association with osteoarthritis. Semin Arthritis Rheum 2012; 42:105-18.
13. Zhang Y, Nevitt M, Niu J, Lewis C, Torner J, Guermazi A, et al. Fluctuation of knee pain and changes in bone marrow lesions, effusions, and synovitis on magnetic resonance imaging. Arthritis Rheum 2011;63:691-9.
14. Crema MD, Roemer FW, Zhu Y, Marra MD, Niu J, Zhang Y, et al. Subchondral cystlike lesions develop longitudinally in areas of bone marrow edema-like lesions in patients with or at risk for knee osteoarthritis: detection with MR imaging—the MOST study. Radiology 2010;256:855-62.
15. Kortekaas MC, Kwok WY, Reijnierse M, Huizinga TW, Kloppenburg M. Follow-up study of inflammatory ultrasound features in hand osteoarthritis over a period of 3 months: variable as well as constant. Osteoarthritis Cartilage 2014;22:40-3.
16. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011;70:1581-6.
17. Bijsterbosch J, Visser W, Kroon HM, Stamm T, Meulenbelt I, Huizinga TW, et al. Thumb base involvement in symptomatic hand osteoarthritis is associated with more pain and functional disability. Ann Rheum Dis 2010;69:585-7.

APPENDIX 1. Definitions and scaling of features included in the OMERACT hand OA MRI score.

	Definition	Scaling
Synovitis[#]	Thickened synovium with enhancement of gadolinium	0 = Normal, 1 = Mild, 2 = Moderate, 3 = Severe
Erosive damage[#]	Subchondral bone loss, including erosions, bone attrition, and bone resorption	0 = Normal; 1 = Mild, ≤ 10% of bone volume or ≤ 25% joint surface affected*; 2 = Moderate, 11–20% of bone volume and/or 26–50% of joint surface affected*; 3 = Severe, > 20% of bone volume and/or > 50% of joint surface affected*
Cyst	A sharply marginated bone lesion with typical signal characteristics* without a cortical break. *Increased signal intensity on T1w FS, T2w FS or STIR images	0 = Normal; 1 = Mild, ≤ 10% of bone volume*; 2 = Moderate, 11–20% of bone volume*; 3 = Severe, > 20% of bone volume*
Osteophyte	Abnormal bone protuberance at joint margins or surfaces	0 = Normal, no osteophytes; 1 = Mild, 1–3 small osteophytes; 2 = Moderate, ≥ 4 small osteophytes and/or ≥ 1 moderate osteophyte(s); 3 = Severe, at least 1 large osteophyte
Cartilage space loss	Loss of the cartilage space based on the interbone distance	0 = Normal; 1 = Mild, loss of cartilage space without bone-to-bone contact; 2 = Moderate, focal complete loss of cartilage space; 3 = Severe, complete cartilage space loss affecting > 50% of the articulating joint area
Malalignment	Angulation or subluxation of joints in the frontal plane	0 = No apparent malalignment; 1 = Mild, ≤ 10° angulation and/or not congruent joint surfaces; 2 = Moderate, 11–20° angulation and/or moderate subluxation without crossing the midline; 3 = Severe, > 20° angulation and/or subluxation with crossing the midline
Bone marrow lesions[#]	A lesion within the trabecular bone with signal characteristics consistent with increased water content (i.e., high signal intensity on STIR images or T2w FS images) and with ill-defined margins.	0 = Normal; 1 = Mild, 1–33% of bone volume*; 2 = Moderate, 34–66% of bone volume*; 3 = Severe, 67–100% of bone volume*

* "Bone volume" and "joint surface" refer to the proximal and distal part of the joint combined. The assessed bone volume should extend from the articular surface to a depth of 0.5 cm for DIP joints and 1.0 cm for PIP joints. [#]0.5 increments for the assessment of synovitis, erosions, and bone marrow lesions. T1w: T1-weighted; T2w: T2-weighted; FS: fat-suppressed; STIR: short-tau inversion recovery; OA: osteoarthritis; MRI: magnetic resonance imaging; DIP: distal interphalangeal; PIP: proximal interphalangeal.