The OMERACT Psoriatic Arthritis Magnetic Resonance Imaging Score (PsAMRIS) Is Reliable and Sensitive to Change: Results from an OMERACT Workshop

PERNILLE BØYESEN, FIONA M. McQUEEN, FRÉDÉRIQUE GANDJBAKHCH, SIRI LILLEGRAVEN, LAURA COATES, CHARLOTTE WIELL, ESPEN A. HAAVARDSHOLM, PHILIP G. CONAGHAN, CHARLES G. PETERFY, PAUL BIRD, and MIKKEL ØSTERGAARD

ABSTRACT. ***Objective.*** The aim of this multireader exercise was to assess the reliability and sensitivity to change of the psoriatic arthritis magnetic resonance imaging score (PsAMRIS) in PsA patients followed for 1 year.

Methods. MRI was acquired from 12 patients with PsA before initiation of treatment and after 12 months. MR images were scored according to PsAMRIS (for synovitis, tenosynovitis, periarticular inflammation, bone marrow edema, bone erosion, and bone proliferation) under standardized conditions, in unknown chronological order. Intraobserver/interobserver reliability was examined by intraclass correlation coefficients (ICC) and sensitivity to change by standardized response means (SRM).

Results. The interobserver reliability of PsAMRIS was high for synovitis, tenosynovitis, periarticular inflammation, and bone edema status and change scores (interobserver ICC 0.87–0.97). The intraobserver reliability was moderate to high (ICC 0.60–0.98) for status and change scores, except for change in periarticular inflammation (ICC 0.33). PsAMRIS sensitivity to change was moderate for synovitis, tenosynovitis, and periarticular inflammation (SRM 0.5–0.8), while poor (SRM 0.1–0.3) for bone marrow edema, erosion, and bone proliferation. Rare occurrence and minimal change contributed to poor SRM and change-score ICC for bone parameters.

Conclusion. This multireader exercise, performed under standardized conditions, confirmed PsAMRIS to have high interobserver and intraobserver reliability for hand PsA. Measures of inflammation were sensitive to change, implying that PsAMRIS may be a valuable tool for monitoring change in inflammation during PsA clinical trials. (J Rheumatol 2011;38:2034–8; doi:10.3899/jrheum.110420)

Key Indexing Terms:
PSORIATIC ARTHRITIS MAGNETIC RESONANCE IMAGING OUTCOME MEASURE

Psoriatic arthritis (PsA) is a heterogeneous disease with diverse presentations and disease outcomes. During the last decade, new biological therapies for PsA have been introduced, with the potential to modify the disease course and suppress both joint inflammation and erosive bone damage. This has led to a need to develop reliable and responsive outcome measures, so that treatment effects can be monitored accurately. Magnetic resonance imaging (MRI) has the advantage of depicting both soft tissue inflammation, affecting articular and periarticular structures, and bone involvement (including bone marrow edema, erosion, and proliferation). In rheumatoid arthritis (RA), the MRI score (RAMRIS) developed by Outcome Measures in Rheumatology Clinical Trials (OMERACT) has been shown to be both reliable and responsive to change in multiple patient groups[1,2,3,4,5]. In 2004, an OMERACT MRI working party began to develop a PsA MRI scoring system (PsAMRIS)[6,7]. This has been shown in previous exercises to demonstrate moderate to good inter-

From the Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway; University of Auckland, Auckland, New Zealand; Department of Rheumatology, Pitiè Salpêtriere Hospital, APHP, Université Paris 6-UPMC, Paris, France; Section of Musculoskeletal Disease, University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK; Department of Rheumatology, Copenhagen University Hospital at Gentofte, Copenhagen, Denmark; Spire Sciences LLC, San Francisco, CA, USA; Division of Medicine, University of NSW, Sydney, Australia; and Department of Rheumatology, Copenhagen University Hospitals at Hvidovre and Glostrup, Copenhagen, Denmark.

P. Bøyesen, MD, PhD, Senior Registrar, Department of Rheumatology, Diakonhjemmet Hospital; F.M. McQueen, MBChB, MD, FRACP, Professor of Rheumatology, Department of Molecular Medicine and Pathology, Faculty of Medicine and Health Sciences, University of Auckland, Auckland, New Zealand; F. Gandjbakhch, MD, Practising Rheumatologist, Department of Rheumatology, Pitiè Salpêtriere Hospital, APHP, Université Paris 6-UPMC; S. Lillegraven, MD, PhD student, Department of Rheumatology, Diakonhjemmet Hospital; L. Coates, MBChB, MRCP, Arthritis Research UK Clinical Research Fellow, Section of Musculoskeletal Disease, University of Leeds; C. Wiell, MD, PhD, Senior Registrar, Department of Rheumatology, Copenhagen University Hospital at Gentofte; E.A. Haavardsholm, MD, PhD, Senior Registrar, Department of Rheumatology, Diakonhjemmet Hospital; P.G. Conaghan, MB, BS, PhD, FRACP, FRCP, Professor of Musculoskeletal Medicine, Section of Musculoskeletal Disease, University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research; C.G. Peterfy, MD, PhD, Chief Executive Officer, Spire Sciences LLC; P. Bird, BMed (Hons), FRACP, PhD, Grad Dip MRI, Senior Lecturer, Division of Medicine, University of NSW; M. Østergaard, MD, PhD, DMSc, Professor of Rheumatology, Department of Rheumatology, Copenhagen University Hospitals at Hvidovre and Glostrup.

Address correspondence to Dr. Bøyesen; E-mail: pernilleboyesen@gmail.com

reader reliability[6,7,8]. The aim of the current multireader exercise was to test inter- and intrareader reliability of PsAMRIS in a group of patients on anti-tumor necrosis factor-α (TNF-α) therapy or standard disease modifying antirheumatic drugs (DMARD) over a 1-year period. Both status and change scores have been recorded under standardized conditions. This allows determination of whether PsAMRIS can be used as an instrument to monitor responses to therapy in PsA.

MATERIALS AND METHODS

Patients. The exercise was performed using MRI scans acquired from 12 patients with PsA (8 from Copenhagen and 4 from Leeds) before the onset of treatment and after 12 months (Table 1). Patient characteristics were as follows: 58% were women, median age was 56.5 years (range 35.0–72.0) and disease duration 4.0 years (range 0.6–11.0). The 8 patients from Copenhagen received anti-TNF-α treatment and the 4 Leeds patients DMARD.

Methods. The MRI scans included images of the metacarpophalangeal (MCP), proximal interphalangeal (PIP), and distal interphalangeal (DIP) joints of the 2nd–5th fingers. The images were acquired using a 0.6 T Philips Panorama MRI unit (Copenhagen) and a 0.2 T Esaote C Scan MRI unit (Leeds). The acquired images from Copenhagen consisted of sagittal and axial STIR sequences, and coronal, axial, and sagittal T1-weighted before and after intravenous administration of gadolinium. From Leeds, T1-weighted axial, sagittal, and coronal images without gadolinium injection were available. For the workshop all images were anonymized and blinded for time sequence; for second reading, images were re-anonymized and re-randomized. FG, FMcQ, MØ, and P. Bøyesen read MR images of the 12 patients twice on identical PACS workstations. The images were scored according to the PsAMRIS system as described[6,8]. In short, for each MCP, PIP, and DIP joint-level synovitis is scored 0–3, tenosynovitis 0–3, periarticular inflammation 0–2, bone marrow edema 0–6, bone proliferation 0–1, and bone erosions 0–20.

Statistics. Baseline and 1-year change of the PsAMRIS features were analyzed as median (minimum, maximum) of the 4 readers' median scores. Changes in PsAMRIS features within patients were tested by Wilcoxon signed-rank tests. Reliability was evaluated by 2-way mixed intraclass correlation coefficients (ICC), smallest detectable difference (SDD), and the smallest detectable change (SDC) for status and change scores, respectively. Average measure ICC (ICCa) was used for interreader reliability and single-measures ICC (ICCs) for intrareader reliability. Bland and Altman's 95% limits of agreement method was used to estimate the SDD and SDC[9,10].

Specifically, we used the random-effects average of the mean residual error of repeated-measures analysis of variance by $\sqrt{2}$ of the number of readers ($\sqrt{k}$). A SDD%max and SDC%max were expressed as the SDD/SDC divided by the maximum observed scores[10]. The PsAMRIS's sensitivity to change was estimated by the standardized response mean (SRM). The SRM was calculated as the mean change divided by the standard deviation of the change score.

RESULTS

The median (range) PsAMRIS results at baseline and 1-year followup are summarized in Table 2. Synovitis, tenosynovitis, and periarticular inflammation were frequently detected in this group of patients with PsA, whereas bone marrow edema and erosions were infrequent. At 1-year followup, tenosynovitis ($p = 0.05$) statistically significantly improved, and there was a numerical, but not statistically significant, improvement in PsAMRIS synovitis and periarticular inflammation. PsAMRIS bone marrow edema, erosions, and bone proliferation did not change during the 1-year followup period.

The inter- and intrareader reliability of PsAMRIS status and change scores is summarized in Table 3. Overall, the PsAMRIS reliability was good to very good, with some exceptions. The 1-year change in PsAMRIS bone proliferation and bone erosions had poor to trivial reliability (ICCa 0.10 and 0.44, respectively). All readers, however, consistently found minimal or no change in these parameters.

The PsAMRIS's sensitivity to change assessed by SRM is summarized in Table 4. Overall, there was moderate responsiveness of synovitis, flexor tenosynovitis, and periarticluar inflammation measured by PsAMRIS (SRM 0.5–0.8). However, bone marrow edema, erosion, and bone proliferation showed poor sensitivity to change in this exercise (SRM 0.1–0.3). PsAMRIS's sensitivity to change during anti-TNF-α therapy is illustrated in Figure 1.

DISCUSSION

In this multireader exercise, the PsAMRIS was tested under standardized conditions with identical radiological workstations and image setup, in addition to the stringent methodology with blinded chronology of scans applied in previous exercises[8]. The previously demonstrated good to very good reliability was reproduced[8]. Further, PsAMRIS was sensitive to improvement in inflammation scores of synovitis, flexor tenosynovitis, and periarticular inflammation during anti-TNF-α therapy.

Table 1. Patients included in the Psoriatic Arthritis Magnetic Resonance Imaging Score (PsAMRIS) workshop.

	Copenhagen, n = 8	Leeds, n = 4	PsAMRIS Features
Treatment	Anti-TNF-α	DMARD	
Assessment timepoints	0 and 12 months	0 and 12 months	
Field strength	0.6 T	0.2 T	
MRI sequences applied			
T1w pre-Gd injection	x	x	Bone erosions, bone proliferation, n = 12
T1w post-Gd injection	x		Synovitis, tenosynovitis, periarticular inflammation, n = 8
STIR	x		Bone marrow edema, n = 8

TNF-α: tumor necrosis factor-α; DMARD: disease-modifying antirheumatic drugs; MRI: magnetic resonance imaging; T1w: t1 weighted MRI sequence; Gd: Gadolinium; STIR: short-tau inversion recovery.

Table 2. Median (minimum, maximum) baseline and 1-year change in PsAMRIS features for overall scores and subjoint regions. All values are given as median (minimum to maximum) of the 4 scorers' median scores.

PsAMRIS Features (range of total score)	Baseline PsAMRIS				One-year Change in PsAMRIS			
	Total score	MCP	PIP	DIP	Total score	MCP	PIP	DIP
Synovitis (0–36)	11.3 (3.5 to 13)	3.0 (1.0 to 4.5)	4.0 (1.5 to .0)	0.75 (0.0 to 3.0)	–1.5 (–11.0 to 3.5)	–1.75 (–5.0 to 1.5)	0.0 (–6.0 to 3.0)	–0.25 (–2.5 to 2.5)
Flexor tenosynovitis (0–36)	5.3 (0.0 to 6.5)	1.25 (0.0 to 5.0)	0.0 (0.0 to 2.0)	0.0 (0.0 to 1.0)	–2.0 (–9.0 to 1.5)	–0.75 (–3.0 to 2.0)	–1.0 (–3.5 to 0.0)	0.0 (–2.0 to 0.0)
Periarticular inflammation (0-24)	0.3 (0.0 to 1.5)	0.0 (0.0 to 0.0)	0.0 (0.0 to 1.5)	0.0 (0.0 to 0.5)	–0.25 (–4.0 to 0.0)	0.0 (–0.5 to 0.0)	–0.25 (–3.5 to 0.0)	0.0 (–1.0 to 0.0)
Bone marrow edema (0–72)	0.0 (0.0 to 1.5)	0.0 (0.0 to 1.0)	0.0 (0.0 to 0.5)	0.0 (0.0 to 0.0)	0.0 (–9.0 to 1.5)	0.0 (–1.5 to 1.0)	0.0 (–4.5 to 0.0)	0.0 (–2.5 to 0.0)
Bone erosion (0–240)	2.0 (0.0 to 6.5)	1.5 (0.0 to 5.0)	0.0 (0.0 to 2.5)	0.0 (0.0 to 0.5)	0.0 (–0.5 to 1.0)	0.0 (0.0 to 1.0)	0.0 (–0.5 to 0.0)	0.0 (0.0 to 0.5)
Bone proliferation (0–12)	3.0 (0.0 to 5.0)	0.0 (0.0 to 1.5)	0.0 (0.0 to 2.0)	0.75 (0.0 to 4.0)	0.0 (–0.5 to 0.5)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)

PsAMRIS: psoriatic arthritis magnetic resonance imaging score; MCP: metacarpophalangeal joint; PIP: proximal interphalangeal joint; DIP: distal interphalangeal joint.

Table 3A. Inter- and intrareader reliability for PsAMRIS status and change scores. Intraclass correlation coefficients. Key to SRM values: Very good, ICC ≥ 0.80; good, 0.50 ≤ ICC < 0.80; poor, 0.20 ≤ ICC < 0.50; trivial, ICC < 0.20.

PsAMRIS Features	Interreader Reliability, ICCa		Intrareader Reliability, ICCs							
			Reader 1		Reader 2		Reader 3		Reader 4	
	Status	Change	Status	Change	Status	Change	Status	Change	Status	Change
Synovitis	0.95	0.95	0.97	0.78	0.73	0.63	0.93	0.65	0.90	0.38
Flexor tenosynovitis	0.97	0.95	0.81	0.92	0.74	0.73	0.93	0.80	0.73	0.90
Bone proliferation	0.77	0.10	0.20	0.26	0.84	0.87	0.79	0.73	0.44	0.46
Periarticular inflammation	0.91	0.91	0.39	NA*	0.95	NA*	0.88	0.33	0.91	NA*
Bone marrow edema	0.87	0.87	0.98	0.63	0.97	0.13	0.99	1.00	0.03	0.80
Bone erosion	0.97	0.44	0.91	0.94	0.93	0.91	0.92	0.91	0.91	0.88

* Not calculable because zero variance. ICCa: average measure intraclass correlation coefficient; ICCs: single measure intraclass correlation coefficient; PsAMRIS: psoriatic arthritis magnetic resonance imaging score.

Table 3B. Inter- and intrareader reliability for PsAMRIS status and change scores. Smallest detectable difference and change.

PsAMRIS Features	Interreader Reliability (SDD/SDC)				Intrareader Reliability (SDD/SDC)*			
	Status Scores		Change Scores		Status Scores		Change Scores	
	SDD	$SDD\%_{max}$	SDD	$SDD\%_{max}$	SDD	$SDD\%_{max}$	SDD	$SDD\%_{max}$
Synovitis	2.78	13.89	3.19	13.28	3.54	27.62	4.71	32.17
Flexor tenosynovitis	3.35	27.91	2.97	19.79	4.71	45.27	4.08	42.10
Periarticular inflammation	2.11	35.18	2.43	26.98	2.17	33.79	2.02	30.65
Bone marrow edema	1.29	9.24	2.00	15.42	1.76	16.07	2.66	41.64
Bone erosion	1.38	19.66	0.84	21.05	1.79	24.44	1.29	42.99
Bone proliferation	2.62	32.71	0.57	18.94	2.74	57.68	0.92	70.08

* Median values of the 4 readers. SDD: smallest detectable differences; SDC: smallest detectable change; PsAMRIS: psoriatic arthritis magnetic resonance imaging score.

In this multireader exercise all PsAMRIS features could be scored in 8 of 12 included patients, all of whom received anti-TNF-α therapy (Copenhagen). PsAMRIS bone erosions and bone proliferation were the only features that could be scored in the remaining 4 DMARD-treated patients (Leeds). As a consequence, all results concerning PsAMRIS synovitis, flexor tenosynovitis, periarticular inflammation, and bone marrow edema are derived from patients treated with anti-TNF-α therapy.

Intrareader reliability for PsAMRIS was examined for the first time in this exercise and was overall "good" to "very good" (0.5 to 1.0). The reliability of PsAMRIS synovitis, bone marrow edema, and erosion scores was comparable with the high reliability for the same components of the RAMRIS

Table 4. Sensitivity to change of the PsAMRIS measured by standardized response means (SRM). Key to SRM values: large, SRM ≥ 0.80; moderate, 0.50 ≤ SRM < 0.80; small, 0.20 ≤ SRM < 0.50; trivial, SRM < 0.20.

PsAMRIS Features	Total Score	Standardized Response Mean MCP Joints	PIP Joints	DIP Joints
Synovitis	–0.50	–0.77	–0.28	–022
Flexor tenosynovitis	–0.79	–0.53	–0.97	–0.57
Periarticular inflammation	–0.54	–0.47	–0.44	–0.72
Bone marrow edema	–0.34	–0.08	–0.46	–0.31
Bone erosion	0.22	0.52	–0.53	0.00
Bone proliferation	–0.11	–0.29	NA*	0.16

* Not calculated because zero variance. PsAMRIS: psoriatic arthritis magnetic resonance imaging score; MCP: metacarpophalangeal joint; PIP: proximal interphalangeal joint; DIP: distal interphalangeal joint.

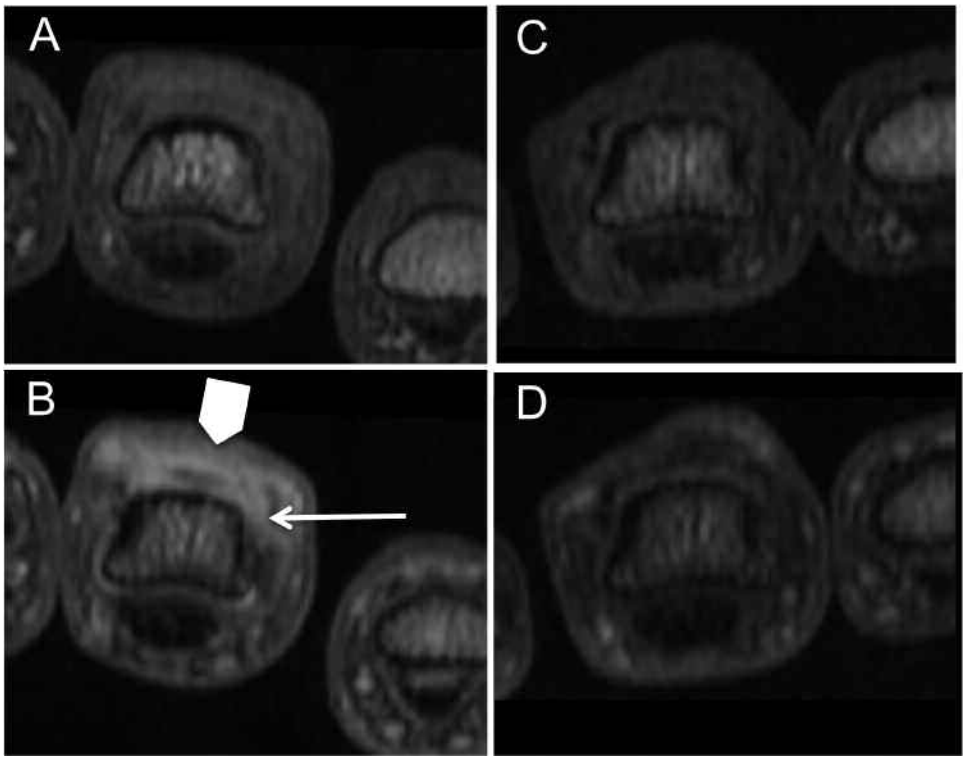

Figure 1. Magnetic resonance imaging examples of PsAMRIS's sensitivity to change during anti-TNF-α treatment. The T1-weighted images show before (A and B) and after (C and D) gadolinium injection. Baseline MR images of PIP-3 (A and B) show periarticular inflammation (broad arrow, B) and synovitis (narrow arrow, B). After 12 months with anti-TNF-α treatment (C and D) the periarticular inflammation and synovitis of the PIP-3 resolved.

when tested in patients with RA[1]. However, the results for the PsA-specific features of periarticular inflammation and bone proliferation were more variable. This may reflect difficulty in readers recognizing and scoring these features, which could be addressed by further training. The intrareader reliability for PsAMRIS change scores was also generally "good" to "very good." Interreader reliability for PsAMRIS status scores was "good" to "very good" and was better than in previous PsAMRIS exercises[8]. This was also true for the reliability of change scores, with the exception of PsAMRIS bone erosion and bone proliferation, where the ICC were reduced due to the lack of change during the 1-year followup, and not to any major disagreements between readers[11]. Other studies in larger and more heterogeneous PsA populations, where more change in bone parameters occurred, would be needed to investigate the responsiveness of these features.

During the 1-year followup tenosynovitis improved and we found a numerical but not statistically significant improvement in PsAMRIS synovitis and periarticular inflammation. The SRM was used to measure responsiveness, as done previously in similar settings[2]. PsAMRIS synovitis, flexor tenosynovitis, and periarticular inflammation were moderately (0.5 to

0.8) sensitive to change, but for bone edema, sensitivity to change was lower (SRM –0.34). A possible explanation of this result was that the bone edema score was high in only one patient, whereas the remaining patients had either mild (score 1) or no bone edema, making it difficult for an improvement to be observed. PsAMRIS bone erosion and bone proliferation were not responsive to change. This was caused by the minimal change that occurred in accordance with the potent influence of anti-TNF-α therapy on bone, not by major disagreements between readers.

A limitation of the study was the small sample size. Further, the external validity of the results could be questioned. The PsA patients studied here all had involvement of the fingers but in patients with, for example, oligoarticular large-joint disease, PsAMRIS may not capture relevant pathology or its change in the setting of a randomized controlled trial. Therefore, the applicability of this scoring system to a more heterogeneous group of PsA patients remains to be examined.

In summary, this multireader exercise performed under standardized conditions has confirmed PsAMRIS to have high interobserver and intraobserver reliability for hand PsA. The MRI measures of inflammation were sensitive to change, implying that PsAMRIS may be a valuable tool for monitoring change in inflammation during PsA clinical trials.

REFERENCES

1. Haavardsholm EA, Østergaard M, Ejbjerg BJ, Kvan NP, Uhlig TA, Lilleås FG, et al. Reliability and sensitivity to change of the OMERACT rheumatoid arthritis magnetic resonance imaging score in a multireader, longitudinal setting. Arthritis Rheum 2005;52:3860-7.
2. Haavardsholm EA, Østergaard M, Hammer HB, Bøyesen P, Boonen A, van der Heijde D, et al. Monitoring anti-TNF-alpha treatment in rheumatoid arthritis: responsiveness of magnetic resonance imaging and ultrasonography of the dominant wrist joint compared with conventional measures of disease activity and structural damage. Ann Rheum Dis 2009;68:1572-9.
3. Ejbjerg BJ, McQueen FM, Lassere MND, Haavardsholm EA, Conaghan PG, O'Connor PJ, et al. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the wrist joint. Ann Rheum Dis 2005;64 Suppl 1:i23-47.
4. Bird P, Conaghan PG, Ejbjerg BJ, McQueen FM, Lassere MND, Peterfy CG, et al. The development of the EULAR-OMERACT rheumatoid arthritis MRI reference image atlas. Ann Rheum Dis 2005;64 Suppl 1:i8-10.
5. Conaghan PG, Bird P, Ejbjerg BJ, O'Connor PJ, Peterfy CG, McQueen FM, et al. The EULAR-OMERACT rheumatoid arthritis MRI reference image atlas: the metacarpophalangeal joints. Ann Rheum Dis 2005;64 Suppl 1:i11-21.
6. Østergaard M, McQueen F, Wiell C, Bird P, Bøyesen P, Ejbjerg B, et al. The OMERACT psoriatic arthritis magnetic resonance imaging scoring system (PsAMRIS): definitions of key pathologies, suggested MRI sequences, and preliminary scoring system for PsA Hands. J Rheumatol 2009;36:1816-24.
7. McQueen FM, Lassere MND, Bird P, Haavardsholm EA, Peterfy CG, Conaghan PG, et al. Developing a magnetic resonance imaging scoring system for peripheral psoriatic arthritis. J Rheumatol 2007;34:859-61.
8. McQueen F, Lassere M, Duer-Jensen A, Wiell C, Conaghan PG, Gandjbakhch F, et al. Testing an OMERACT MRI scoring system for peripheral psoriatic arthritis in cross-sectional and longitudinal settings. J Rheumatol 2009;36:1811-5.
9. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307-10.
10. Lassere M, Boers M, van der Heijde D, Boonen A, Edmonds J, Saudan A, et al. Smallest detectable difference in radiological progression. J Rheumatol 1999;26:731-9.
11. Lassere MND, McQueen FM, Østergaard M, Conaghan PG, Shnier R, Peterfy CG, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Exercise 3: an international multicenter reliability study using the RA-MRI Score. J Rheumatol 2003;30:1366-75.