Developing a Magnetic Resonance Imaging Scoring System for Peripheral Psoriatic Arthritis

FIONA McQUEEN, MARISSA LASSERE, PAUL BIRD, ESPEN A. HAAVARDSHOLM, CHARLES PETERFY, PHILIP G. CONAGHAN, BO EJBJERG, HARRY GENANT, PHILIP O'CONNOR, PAUL EMERY, and MIKKEL ØSTERGAARD

ABSTRACT. We describe the first steps in developing an OMERACT magnetic resonance imaging (MRI) scoring system for peripheral psoriatic arthritis (PsA). A preexisting MRI dataset (finger joints) from 10 patients with PsA was scored by 4 readers for bone erosion, bone edema, synovitis, tendinopathy, and extracapsular features of inflammation (including enthesitis) according to specified criteria. Scoring reliability between readers was moderate to high for bone edema and erosion, but lower for soft tissue inflammation. Measures to improve reliability for future exercises will include reviewing definitions of pathological features and prior reader calibration. (J Rheumatol 2007;34:859–61)

Key Indexing Terms:
PSORIATIC ARTHRITIS SYNOVITIS BONE EROSIONS BONE EDEMA
TENDINOPATHY EXTRACAPSULAR INFLAMMATION RELIABILITY

Psoriatic arthritis (PsA) has a distinctive set of radiographic features that help differentiate this condition from rheumatoid arthritis (RA) and other inflammatory arthropathies[1]. The literature concerning the magnetic resonance imaging (MRI) features of PsA is sparse[2,3]. While bone erosions and synovitis seem similar to their equivalents in RA [although the distal interphalangeal (DIP) joints may be involved], other spondyloarthropathic features such as dactylitis, periostitis, and enthesitis have been less well described and provide an additional layer of complexity. With the advent of highly effective biologic therapies[4], it has recently become possible to influence disease progression in PsA, and there is now a need to quantify changes in articular inflammation and damage so that the efficacy of such new treatments can be accurately assessed. The RA MRI score (RAMRIS) was developed through an OMERACT iterative process to capture disease progression in RA[5]. There are currently no validated scoring systems for MRI in peripheral PsA, so we propose to develop one using the same procedure, acknowledging that there will be constraints imposed by the imaging modality itself and the data sets available.

In this initial exercise, we scored images from a preexisting PsA MRI dataset, using a system based on a RAMRIS framework but with additional categories to include PsA-specific features such as extracapsular inflammation. The aim of our project was to determine the pathological features with the greatest interreader reliability for inclusion in a preliminary PsA MRI score (PAMRIS).

MATERIALS AND METHODS

The OMERACT MRI group began development of PAMRIS after consensus meetings at American College of Rheumatology 2004 and European League Against Rheumatism 2005. Synovitis and bone erosions were scored as in the RAMRIS system (0–3 and 0–10, respectively). Bone edema was scored 0–3 and in addition, categorized as subchondral, entheseal, or diaphyseal. Extracapsular inflammation was scored as absent or present (0 or 1). Tendinopathy was also evaluated: tenosynovitis (0–3), intratendinous edema/enhancement (0–1), and edema/enhancement at insertion (0–1). While it was recognized that other features such as periostitis, bony proliferation, and ankylosis might also be present on some scans, no attempt was made to score these in this first exercise as they were very infrequent.

An image set of MRI scans from 10 PsA patients was chosen (by Charlotte Wiell, Copenhagen). These included images of the 2nd-5th fingers [MCP, proximal interphalangeal (PIP), and DIP joints] obtained on a 0.6 T Philips Panorama MRI unit using the following sequences: 3-D T1 weighted

From the Department of Rheumatology, Auckland University, Auckland, New Zealand; Department of Rheumatology, St. George Hospital, University of New South Wales, Sydney, Australia; Diakonhjemmet Hospital, University of Oslo, Oslo, Norway; Synarc Inc., University of California, San Francisco, California, USA; Academic Unit of Musculoskeletal Disease, University of Leeds; Department of Radiology, Chapel Allerton Hospital, Leeds, UK; and Copenhagen University Hospitals at Herlev and Hvidovre, Copenhagen, Denmark.

Travel for Dr. McQueen to attend OMERACT 8 in Malta was assisted by grants awarded by the Auckland Medical Research Foundation and the Maurice and Phyllis Paykel Trust.

F. McQueen, MD, FRACP, Associate Professor in Rheumatology, Department of Rheumatology, Auckland University; M. Lassere, MB, BS, Grad Dip Epi, PhD, FRACP, FAFPHM, Associate Professor in Medicine, Department of Rheumatology, St. George Hospital, University of NSW; P. Bird, BMed (Hons) FRACP, PhD, Grad Dip MRI, Senior Lecturer, University of NSW; E.A. Haavardsholm, MD, Research Fellow, Diakonhjemmet Hospital, University of Oslo; C. Peterfy, MD, PhD, Chief Medical Officer, Synarc Inc.; P.G. Conaghan, MB BS, PhD, FRACP, FRCP, Professor of Musculoskeletal Medicine, Academic Unit of Musculoskeletal Disease, University of Leeds; B. Ejbjerg, MD, PhD, Senior Registrar, Copenhagen University Hospital at Herlev; H. Genant, MD, FACR, FRCR, Professor of Radiology, University of California, San Francisco; P. O'Connor, MB BS, MRCP, FRCR, Consultant Skeletal Radiologist, Department of Radiology, Chapel Allerton Hospital; P. Emery, MA, MD, FRCP, ARC Professor in Rheumatology, Academic Unit of Musculoskeletal Disease, University of Leeds; M. Østergaard, MD, PhD, DMSc, Professor in Rheumatology/Arthritis, Copenhagen University Hospitals at Herlev and Hvidovre.

Address reprint requests to Dr. F. McQueen, Department of Molecular Medicine and Pathology, University of Auckland, PO Box 92019, Auckland, New Zealand. E-mail: f.mcqueen@auckland.ac.nz

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

(T1w) gradient-echo (0.5 × 0.5 × 1.0 mm) with subsequent axial, coronal, and sagittal reconstruction before and after intravenous contrast injection with 0.1 mmol gadolinium–DTPA BMA/kg body weight (Omniscan, Amersham Health) plus sagittal short-tau inversion recovery (STIR) images. Scans were circulated on CD-ROM to 4 readers (EH, FM, PB, MØ) and were read using the commercially available software package Merge eFilm. After initial category definitions had been agreed, there was no further attempt at reader calibration prior to scoring. Analyses of variation were performed for all scores to determine whether there was any significant difference between readers plus post-hoc tests using Tukey HSD to identify the different reader.

Interreader single- and average-measure intraclass correlation coefficients (ICC)[6] were calculated across all joints for all measures scored, and readers were asked to record specific comments regarding areas of difficulty. ICC are a relative measure of agreement. The single-measure fixed effects ICC is similar to the quadratic weighted kappa for ordinal scale measures, where the weighted kappa is agreement beyond chance agreement. The average measure ICC corrects for the number of readers so this was also provided.

RESULTS

Interreader reliability. When mean scores were compared between the 4 readers, there were no significant differences for synovitis, bone edema, tenosynovitis, or extracapsular inflammation. However, Reader 1 scored higher than the others for bone erosions ($p < 0.001$; Table 1). When this reader was excluded, there was no significant difference between scores for the other readers ($p = 0.32$). Data for reliability of erosion scores were therefore analyzed both as 3-reader and 4-reader scores.

Table 2 shows single-measures interreader ICC for all components of the score. While the 3-reader ICC for bone erosion (0.91) was very good, and for bone edema was moderate (0.63), interreader reliability for synovitis, tenosynovitis, and extracapsular inflammation was low.

Difficulties for readers. Several readers recorded difficulty in assessing damage and inflammation at the very small joints (PIP and DIP) where image resolution was sometimes poor. It was also felt that recording patterns of bone edema was too difficult, as there was little to separate "subchondral" from "entheseal" in many cases and "diaphyseal" was very rare. Readers also recorded problems defining the exact location of entheseal regions adjacent to small finger joints making the category "extracapsular inflammation" difficult to assess. This was also felt to overlap with the category "edema/enhancement at tendon insertion." When synovitis coexisted with entheseal and other extracapsular inflammation, with all areas showing increased signal on STIR and post-Gd T1 weighted images, it was also difficult to allocate separate scores for each at the small joints.

Table 2. Interreader single and average measure ICC for sum of scores across all joints (MCP, PIP, and DIP joints).

	Single Measure ICC	Average Measure ICC
Bone erosion		
3-rater	0.91	0.97
4-rater	0.57	0.84
Bone edema	0.63	0.86
Synovitis	0.21	0.53
Tenosynovitis	0.19	0.49
Extracapsular inflammation	0.07	0.25

While there were examples of florid flexor tenosynovitis within the image set that were recognized by all readers (Figure 1), overall scoring of tendinopathy was only fair, possibly because of difficulties differentiating periostitis from tendinopathy on sagittal STIR images.

DISCUSSION

This preliminary multireader exercise has indicated that creating an MRI scoring system for PsA is possible, with moderate to very good interreader reliability for bone inflammation and damage features (bone edema and erosion). However, difficulties were encountered in assessing synovitis, especially at the very small PIP and DIP joints, and this was apparent from lower interreader reliability than has been achieved using the RAMRIS system in RA[7]. Reliability was also low for scoring extracapsular inflammation and tendinopathy. These difficulties were expected, as clear definitions for the new PsA pathologies had not been provided to readers and there was no pre-exercise reader calibration to optimize their recognition. The scoring sheet used was deliberately overinclusive to capture a wide range of pathologies, with the intention that it could be refined later to incorporate only those features that could be reliably recognized and scored. This may sometimes mean omitting features that are important pathologically but are not well seen on peripheral MRI, as was the case for cartilage involvement, which was omitted during development of RAMRIS[8].

Another aspect of designing an MRI scoring system for PsA is deciding on the optimal sequences and acquisitions to

Table 1. Scores* (mean and range) for synovitis, bone edema, bone erosions and extracapsular signs of inflammation from 10 PsA MRI scans of the fingers (MCP, PIP, and DIP joints).

	Synovitis		Bone Edema		Bone Erosion		Extracapsular Inflammation	
	Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max	Mean (SD)	Min-Max
Reader 1	8.1 (6)	0–16	1.0 (1.9)	0–16	12.7* (4.5)	10–20	1.4 (2.1)	0–7
Reader 2	6.2 (5)	1–17	2.4 (2.9)	1–17	4.7 (4.6)	1–15	0.6 (0.7)	1–2
Reader 3	6.7 (5.4)	0–18	0.9 (1.5)	0–18	2.5 (3.4)	0–9	0.8 (1.1)	0–3
Reader 4	7.1 (5.3)	0–15	1.2 (2.6)	0–15	2.0 (3.2)	0–9	1.2 (1.5)	0–4

* No significant differences in scores between readers except for erosions, where Reader 1 scores were higher ($p < 0.001$).

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.

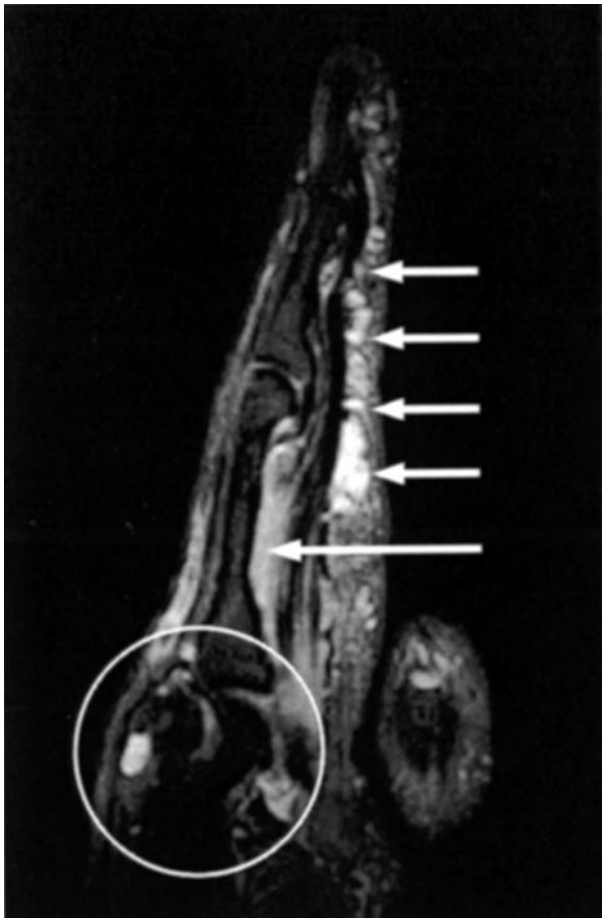

Figure 1. Sagittal STIR image from a patient with PsA showing flexor tenosynovitis (scored at a maximum of 3 by all readers; long arrow) and soft tissue edema of the index finger (short arrows) indicating dactylitis. There is increased signal adjacent to the 2nd metacarpal head, indicating synovitis at the MCP joint (circle).

use. This is more complex than in RA for a number of reasons. There are a broader range of tissues and sites of potential pathology in PsA, but at the same time often fewer joints involved per patient, and these may be asymmetrically distributed. When deciding which pulse sequences to use, spatial resolution and signal-to-noise ratio issues need consideration. T1 weighted SE sequences with fat suppression, pre- and post-contrast with IV gadolinium, are appropriate for most bony lesions including erosions (and were used in this exercise). However, experience in ankylosing spondylitis has suggested that axial T1 weighted sequences lack sensitivity for detection of syndesmophytes[9] and this might also apply to imaging the proliferative bony lesions of peripheral PsA. STIR and T2 weighted fast spin-echo fat-suppressed sequences are very effective for imaging bone edema and soft tissue inflammation[10], but T2 weighted images are time-consuming to acquire and therefore susceptible to movement artefacts. In this exercise, sagittal STIR sequences of the fingers were useful in identifying dactylitis, usually due to tenosynovitis, as they captured images of the complete ray from the MCP joint to the fingertip. In some cases, false positives could have been scored for tenosynovitis due to periostitis causing increased signal on the inner aspect of the tendon sheath.

To develop PAMRIS further we now intend to focus on the "poorly recognized" categories identified in this exercise, such as extracapsular inflammation and tendinopathy, and decide on definitions, both in terms of anatomic localization and typical MRI signal characteristics. Agreement on image acquisition, sequences, and planes of imaging needs to be reached and future exercises should include pre-exercise reader calibration. From this beginning we hope to refine the process and eventually develop a scoring system that conforms to the OMERACT principles of truth, discrimination, and feasibility.

REFERENCES

1. Gladman DD. Psoriatic arthritis. In: Khan MA, editor. Ankylosing spondylitis and related spondyloarthropathies. Spine: State of the art review. Philadelphia: Hanley and Belfus Inc; 1990:637-56.
2. McQueen F, Lassere M, Ostergaard M. Magnetic resonance imaging in psoriatic arthritis: a review of the literature. Arthritis Res Ther 2006;8:207. Epub 2006 Mar 23.
3. Ory PA, Gladman DD, Mease PJ. Psoriatic arthritis and imaging. Ann Rheum Dis 2005;64 Suppl II:ii55-ii57.
4. Mease PJ. Adalimumab: an anti-TNF agent for the treatment of psoriatic arthritis. Expert Opin Biol Ther 2005;5:1491-504.
5. Ostergaard M, Peterfy C, Conaghan P, 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;30:1385-8.
6. Shrout P, Fleiss J. Intra-class correlations: using in assessing rater reliability. Psychol Bull 1979;86:420–8.
7. Haavardsholm EA, Østergaard M, Ejbjerg B, 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.
8. McQueen FM, Østergaard M, Conaghan P, et al. OMERACT rheumatoid arthritis magnetic resonance imaging studies. Summary of OMERACT 6 MR imaging module. J Rheumatol 2003;30:1387-92.
9. Braun J, Baraliakos X, Golder W, et al. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab. Arthritis Rheum 2003;48:1126-36.
10. Roemer FW, Guermazi A, Lynch JA, et al. Short tau inversion recovery and proton density-weighted fat suppressed sequences for the evaluation of osteoarthritis of the knee with a 1.0 T dedicated extremity MRI: development of a time-efficient sequence protocol. Eur Radiol 2005;15:978-87.

Personal non-commercial use only. The Journal of Rheumatology Copyright © 2007. All rights reserved.