Significance of Clinical Evaluation of the Metacarpophalangeal Joint in Relation to Synovial/Bone Pathology in Rheumatoid and Psoriatic Arthritis Detected by Magnetic Resonance Imaging

MILLICENT A. STONE, LAWRENCE M. WHITE, DAFNA D. GLADMAN, ROBERT D. INMAN, SAM CHAYA, MATTHEW LAX, DAVID SALONEN, DEBORAH A. WEBER, JUDY A. GUTHRIE, EMMA POMEROY, DOMINIK PODBIELSKI, and EDWARD C. KEYSTONE

ABSTRACT. ***Objective.*** Rheumatologists base many clinical decisions regarding the management of inflammatory joint diseases on joint counts performed at clinic. We investigated the reliability and accuracy of physically examining the metacarpophalangeal (MCP) joints to detect inflammatory synovitis using magnetic resonance imaging (MRI) as the gold standard.

Methods. MCP joints 2 to 5 in both hands of 5 patients with rheumatoid arthritis (RA) and 5 with psoriatic arthritis (PsA) were assessed by 5 independent examiners for joint-line swelling (visually and by palpation); joint-line tenderness by palpation (tender joint count, TJC) and stress pain; and by MRI (1.5 Tesla superconducting magnet). Interrater reliability was assessed using kappa statistics, and agreement between examination and corresponding MRI assessment was assessed by Fisher's exact tests ($p < 0.05$ considered statistically significant).

Results. Interrater agreement was highest for visual assessment of swelling ($\kappa = 0.55$–0.63), slight-fair for assessment of swelling by palpation ($\kappa = 0.19$–0.41), and moderate ($\kappa = 0.41$–0.58) for assessment of joint tenderness. In patients with RA, TJC, stress pain, and visual swelling assessment were strongly associated with MRI evaluation of synovitis. Visual swelling assessment demonstrated high specificity (> 0.8) and positive predictive value (= 0.8). For PsA, significant associations exist between TJC and MRI synovitis scores ($p < 0.01$) and stress pain and MRI edema scores ($p < 0.04$). Assessment of swelling by palpation was not significantly associated with synovitis or edema as determined by MRI in RA or PsA ($p = 0.54$–1.0).

Conclusion. In inflammatory arthritis, disease activity in MCP joints can be reliably assessed at the bedside by examining for joint-line tenderness (TJC) and visual inspection for swelling. Clinical assessment may have to be complemented by other methods for evaluating disease activity in the joint, such as MRI, particularly in patients with PsA. (First Release Nov 15 2009; J Rheumatol 2009;36:2751–7; doi:10.3899/jrheum.080205)

Key Indexing Terms:
RHEUMATOID ARTHRITIS PSORIATIC ARTHRITIS PHYSICAL EXAMINATION
MAGNETIC RESONANCE IMAGING REPRODUCIBILITY OF RESULTS

From the University of Toronto; Mount Sinai Hospital; Toronto Western Hospital; and St. Michael's Hospital, Mobility Program Clinical Research Unit, Toronto, Ontario, Canada; The Royal National Hospital for Rheumatic Diseases; and the University of Bath, Bath, UK.

Supported by an educational grant from Centocor Inc./Schering Plough.

M.A. Stone, MD, University of Toronto, St. Michael's Hospital, Mobility Program Clinical Research Unit, The Royal National Hospital for Rheumatic Diseases, and University of Bath; L.M. White, MD, University of Toronto and Mount Sinai Hospital; D.D. Gladman, MD; R.D. Inman, MD, University of Toronto and Toronto Western Hospital; S. Chaya, MD, University of Toronto; M. Lax, MD; D. Salonen, MD, University of Toronto, Mount Sinai Hospital, and Toronto Western Hospital; D.A. Weber, RN; J.A. Guthrie, RN, Mount Sinai Hospital; E. Pomeroy, MA, The Royal National Hospital for Rheumatic Diseases; D. Podbielski, MD, St. Michael's Hospital; E.C. Keystone, MD, University of Toronto and Mount Sinai Hospital.

Address correspondence to Dr. E. Keystone, Rebecca McDonald Centre for Arthritis and Autoimmune Diseases, Room 2-006, 60 Murray St., Toronto, Ontario M5G 1X5, Canada. E-mail: ekeystone@mtsinai.on.ca

Accepted for publication May 12, 2009.

Rheumatologists base many of their clinical decisions regarding the management of inflammatory joint diseases such as rheumatoid arthritis (RA) and psoriatic arthritis (PsA) on joint counts (tender and swollen joints) performed in the clinic. The number of actively inflamed joints at initial presentation predicts subsequent performance and function of patients with recent-onset inflammatory synovitis[1,2]. Joint swelling is particularly important since it is one of several prognostic outcomes that correlates with radiographic progression in both RA and PsA[3,4].

Boers, *et al* demonstrated a direct correlation between swelling in an individual joint and radiographic progression in that joint[1]. Tender joints correlate less well with radiographic progression[5]. Therefore, early detection of synovitis and aggressive treatment could prevent further and longterm disability.

One of the issues facing clinicians is the relevance of detecting joint tenderness as opposed to joint swelling in relation to synovitis. Whether joint tenderness in the absence of joint swelling indicates a significant degree of synovitis remains unclear.

The ability to detect synovitis by clinical examination may differ with different forms of arthritis. The inflamed metacarpophalangeal (MCP) joint in patients with RA is usually soft and described as "boggy." The MCP joints of patients with PsA generally are less tender with less obvious synovitis[6]. The inflamed PsA joint appears more firm or thickened rather than "spongy." The character of the inflamed MCP joint in PsA may result from increased extrasynovial rather than intrasynovial inflammation and proliferative synovitis seen in RA[7].

The purpose of our study was to determine the reliability of physical examination of the MCP joint to detect synovitis using magnetic resonance imaging (MRI) as a gold standard. The assessment of MCP joints was chosen since they are frequently involved early and often undergo the earliest radiographic changes with subsequent deformity and disability. The study was specifically designed to determine the significance of joint tenderness versus swelling with regard to the correlation with synovitis, the sensitivity of techniques to detect joint swelling, and differential sensitivity of detecting synovitis in patients with RA versus PsA. We hypothesized that (1) joint swelling, not joint tenderness, would more closely correlate with intrasynovial synovitis; and (2) physical examination would be more reliable in detecting synovitis in RA than in PsA.

MATERIALS AND METHODS

Patients. Five examiners independently assessed swelling and tenderness in the 2nd to 5th MCP joints of 10 patients with inflammatory arthritis. Five patients had established RA, and 5 had PsA. Patients were required to have at least 2 swollen MCP joints (1 swollen MCP joint in each hand) and 1 of these had to be classified as having grade 1 swelling (defined as swelling altering the joint-line margin as noted on inspection) to be eligible for inclusion. Patients were screened for suitability 1 week prior to the study and no alteration in medications was permitted within a week of commencement of the study. An independent health professional, not one of the examiners, selected the patients for inclusion based on these prespecified criteria. All patients signed written informed consent and the study was approved by the Institutional Review Boards at the University Health Network and Mount Sinai Hospital.

Physical examination maneuvers. The 5 examiners were experienced in the examination of MCP joints. A demonstration session was conducted before the assessments where all examiners agreed by consensus on the methods of assessment. Examiners were randomized as to their order of evaluation. Joint-line tenderness and swelling were evaluated at the same time. Examiners were blinded to patient details. Although every effort was made to blind examiners to diagnosis, examiners were experienced rheumatologists and in certain patients it was clinically apparent which form of inflammatory arthritis was present. The same examiners were used for the examination of RA and PsA patients.

The 2nd to 5th MCP joints were evaluated for joint swelling by the standard 2-finger or newly developed 4-finger technique, as described below. A joint swelling score was performed as a modification of the technique of Thompson, *et al* on a 4 point Likert scale (0: no swelling, 3: maximum swelling)[8]. Each technique was conducted on all 4 MCP joints at different times as described below. The order of examinations by an individual examiner using each technique was prespecified by a computer-generated randomization grid in order to reduce the systemic influence of one technique or another.

In the standard 2-finger technique the finger is held in slight flexion and the joint is palpated with the thumbs on the dorsal aspect of the finger on either side of the extensor tendons. In the 4-finger technique the finger is held in slight extension, and the examiner's thumb and index finger of one hand surround the distal aspect of the metacarpal head at the joint, while the thumb and index finger of the other hand surround the proximal aspect of the metacarpal head proximally (Figure 1). The capturing fingers ballot for spongy swelling using a graded subjective score ranging from 0 (no swelling) to 3 (severe swelling). Swelling was also assessed by visual inspection from the dorsal aspect of the hand with the MCP joint held in neutral position (0 = no swelling, 1 = unable to clearly identify extensor tendons, 2 = swelling around joint line, 3 = pannus and obvious visible swelling around joint margin). Visual inspection at rest was also recorded using a dichotomous assessment (swelling/no swelling).

Tenderness of the MCP joint (tender joint count, TJC) was evaluated by applying pressure over the joint line using a graded score ranging from 0 (no tenderness) to 3 (wincing and withdrawing). Stress pain was defined as pain elicited at the extreme of passive extension and was recorded as a dichotomous variable (present/absent).

Imaging methods. All 10 patients were imaged using a twin-speed 1.5 Tesla MR scanner (Sigma, General Electric Healthcare, Milwaukee, WI, USA) within 24 h of the clinical examination. Patients were positioned prone on the scanner table with the arm fully abducted and extended over the head and the hand to be imaged placed as closely as possible to the isocenter of the scanner. Phase-coupled dual 3 inch ring coils placed over the dorsal and volar aspects of the MCP joint were used for imaging. The field of view was 10 x 10 cm and was centered on the 2nd to 5th MCP joint. The imaging protocol comprised axial and coronal spin-echo T1-weighted and fast spin-echo fat-suppressed T2-weighted sequences, followed by axial and coronal fat-suppressed T1-weighted acquisitions obtained immediately (less than 2 min) following intravenous injection of 0.1 mmol/kg body weight of gadolinium contrast material (Omniscan, Amersham Health, Princeton, NJ, USA). All the imaging sequences used a slice thickness of 3 mm and 0 mm gap.

The Outcome Measures in Rheumatology RAMRIS Version 3 scoring system was employed. The extent of bone erosion, bone defects, bone edema, and synovitis was evaluated[9,10] (see Figure 2).

Bone erosions, bone defects, and bone edema were measured at 8 sites, namely, the base of the proximal phalanx and the head of the metacarpal at each of the 2nd to 5th MCP joints. Therefore, the score for each MCP joint was 0 to 20, and the aggregated score for all 4 joints was 0 to 80, where a score of 0 indicated no erosion and a score of 80 indicated no bone. Bone defects and edema were scored similarly.

Synovitis was determined by gadolinium enhancement of the synovial compartment. It was assigned a score of 0 to 3, where 0 was normal with no synovial enhancement or enhancement no thicker than the joint capsule. Scores of 1 to 3 (mild, moderate, severe) increased by thirds of the presumed maximum volume of enhancing tissue in the synovial compartment. This global score was assigned to the 4 MCP joints, giving an aggregated score of 0 to 12.

Statistical methods. The degree of agreement (interrater reliability) of examination measurements per diagnosis was assessed using kappa (κ) statistics. These statistics take on a maximum value of 1, high positive values indicating a strong degree of agreement among examiners. Following the guidelines suggested by Kuritz, *et al*[11], a κ statistic above 0.80 is considered to demonstrate near-perfect reliability, 0.61–0.80 substantial agreement, 0.41–0.60 moderate agreement, 0.21–0.40 fair agreement, 0.00–0.20 slight agreement, and < 0.00 is an indication of poor reliability of measurements taken by different examiners on the same joint.

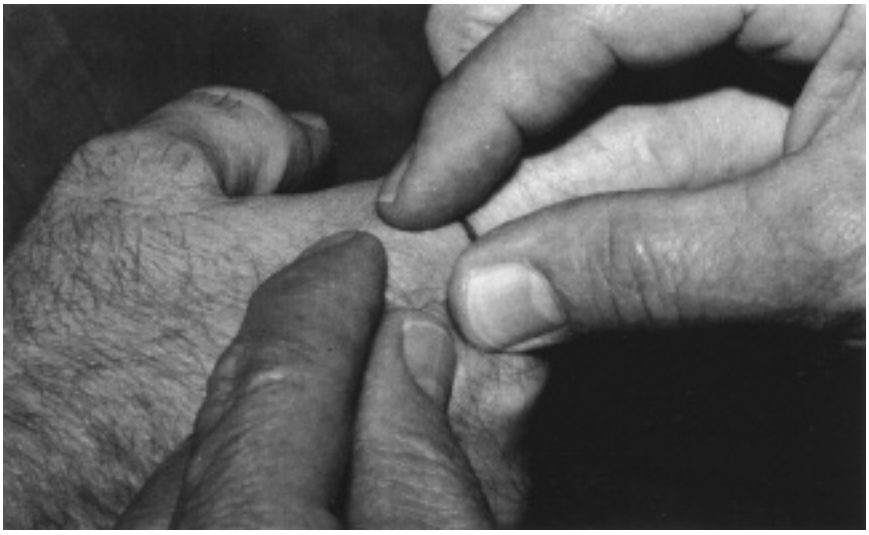

Figure 1. Four-finger technique for assessment of joint-line swelling. The finger is held in slight flexion, and the examiner's 1st and 2nd fingers of the same hand surround the joint distally on either side of the extensor tendon, while the 1st and 2nd fingers of the other hand surround the joint proximally on either side of the extensor tendon. The capturing fingers ballot for swelling and effusion using a graded subjective score ranging from 0 (no swelling) to 3 (severe swelling).

To assess the degree of agreement between examination assessments (as measured by the 2-finger technique, 4-finger technique, stress, TJC grading, and visual at-rest grading and classification techniques) and MRI assessments (which are broken down into synovitis, erosion, edema, and defect subscales), a series of Fisher's exact tests were performed (these tests were used instead of chi-squared tests due to the relatively small sample sizes in some cases). All clinical examination measures and MRI assessments were dichotomized as 0 = outcome absent versus 1 = outcome present. We dichotomized them rather than using the graded scaling when comparing to MRI. These tests were performed separately for RA and PsA subjects and for examination measurements provided by each of the raters. Since each rater provided 8 examination assessments per patient (for each of 8 joints), the individual joint assessments were compared to the MRI assessments for the corresponding joint.

Statistical significance was defined by p value < 0.0004 based on the Bonferroni correction for multiple comparisons. All statistical analysis was carried out using SAS System v.8.2.

RESULTS

Interrater reliability for physical examination maneuvers. Joint swelling: There was fair agreement among examiners for assessment of joint swelling using the 2-finger technique [$\kappa = 0.27$, 95% confidence interval (CI) = 0.20–0.34] and 4-finger technique ($\kappa = 0.30$, 95% CI 0.23–0.37). Visual assessment of swelling at rest demonstrated the highest degree of agreement among examiners, with a κ value of 0.61 (95% CI 0.54–0.68). When agreement was analyzed by disease subtype there was only slight agreement among examiners in the assessment of swelling by palpation in PsA (2-finger technique: $\kappa = 0.19$, 95% CI 0.09–0.29; 4-finger technique: $\kappa = 0.13$, 95% CI 0.03–0.23) as compared to fair agreement when examiners evaluated patients with RA (2-finger technique: $\kappa = 0.34$, 95% CI 0.24–0.44; 4-finger technique: $\kappa = 0.41$, 95% CI 0.31–0.51; Table 1). Visual assessment of swelling exhibited substantial agreement in RA ($\kappa = 0.63$, 95% CI 0.53–0.73, $p < 0.001$) and modest agreement among examiners in PsA ($\kappa = 0.55$, 95% CI 0.45–0.65, $p < 0.001$).

Joint tenderness: Overall assessment of joint tenderness (by stress pain and joint-line tenderness assessments) demonstrated a moderate degree of agreement, with κ values of 0.41–0.58. Interestingly, there was better agreement among examiners in the assessment of tenderness in patients with PsA compared to patients with RA (Table 1).

Comparison of physical examinations to MRI for assessment of joint activity by disease subtype. As noted above, all clinical examination measures and MRI assessments were dichotomized as 0 = outcome absent versus 1 = outcome present.

RA: Among patients with RA, joint-line tenderness, stress pain, and visual assessment of swelling correlated with MRI evaluation of synovitis ($p < 0.01$, $p < 0.01$, and $p = 0.02$, respectively). There was no correlation between swelling as determined by palpation assessments (2-finger or 4-finger) and MRI scores (Table 2). Visual assessment of swelling demonstrated high specificity (> 0.8) and positive predictive value (PPV) (0.8) for detection of synovitis in the joint. In addition, stress pain was associated with MRI edema scores ($p = 0.02$) and exhibited a high PPV of 0.81 (Table 2).

PsA: For PsA, an association between TJC and MRI synovitis scores ($p < 0.01$) and a high negative predictive value (NPV) of 0.82 was noted (Table 2). In addition, there was an association between stress pain and MRI edema scores ($p < 0.04$). However, no association was demonstrated between

Table 1. Interrater reliability, by disease subtype, among examiners with physical evaluation techniques.

Disease Subtype	Evaluation Technique	Kappa (95% CI)	Significance
RA	2–finger technique	0.34 (0.24, 0.44)	< 0.0001
	4-finger technique	0.41 (0.31, 0.51)	< 0.0001
	Stress	0.31 (0.23, 0.39)	< 0.0001
	TJC	0.22 (0.14, 0.28)	< 0.0001
	Visual at-rest Y/N	0.63 (0.53, 0.73)	< 0.0001
	Visual at-rest grade	0.40 (0.34, 0.46)	< 0.0001
PsA	2-finger technique	0.19 (0.09, 0.29)	< 0.0001
	4-finger technique	0.13 (0.03, 0.23)	0.0061
	Stress	0.50 (0.43, 0.57)	< 0.0001
	TJC	0.58 (0.50, 0.66)	< 0.0001
	Visual at-rest Y/N	0.55 (0.45, 0.65)	< 0.0001
	Visual at-rest grade	0.45 (0.37, 0.53)	< 0.0001

RA: rheumatoid arthritis; PsA: psoriatic arthritis; TJC: tender joint count.

Table 2. Association of physical examination assessment of MCP joints (dichotomized) with MRI synovitis, erosion, and edema scores (dichotomized) in subjects with RA and PsA.

		MRI Measurement					
		RA			PsA		
Examination Measurement	Statistic	Synovitis	Erosion	Edema	Synovitis	Erosion	Edema
2-finger	p	0.64	0.32	0.54	0.53	0.53	0.68
	Sensitivity	0.44	0.44	0.53	0.49	0.49	0.48
	Specificity	0.63	0.44	0.55	0.58	0.58	0.58
	PPV	0.67	0.44	0.64	0.39	0.39	0.43
	NPV	0.40	0.44	0.44	0.68	0.68	0.63
4-finger	p	0.65	0.32	0.54	0.39	0.67	1.0
	Sensitivity	0.46	0.46	0.55	0.43	0.40	0.38
	Specificity	0.60	0.42	0.53	0.66	0.65	0.63
	PPV	0.66	0.44	0.63	0.41	0.38	0.41
	NPV	0.40	0.44	0.44	0.68	0.67	0.60
Stress	p	< 0.01	0.82	0.02	0.41	0.01	0.04
	Sensitivity	0.28	0.28	0.55	0.54	0.26	0.35
	Specificity	0.97	0.76	0.88	0.55	0.40	0.43
	PPV	0.93	0.54	0.81	0.40	0.19	0.29
	NPV	0.45	0.51	0.47	0.69	0.50	0.50
TJC	p	< 0.01	0.03	0.08	< 0.01	0.83	0.84
	Sensitivity	0.32	0.32	0.28	0.74	0.49	0.53
	Specificity	1.0	0.88	0.88	0.62	0.48	0.50
	PPV	1.0	0.73	0.77	0.51	0.33	0.41
	NPV	0.47	0.56	0.45	0.82	0.63	0.61
Visual at-rest	p	0.02	0.84	0.07	0.40	0.83	0.5
	Sensitivity	0.48	0.48	0.58	0.49	0.40	0.38
	Specificity	0.80	0.48	0.63	0.62	0.57	0.55
	PPV	0.80	0.48	0.70	0.40	0.33	0.36
	NPV	0.48	0.48	0.50	0.69	0.64	0.57

MCP: metacarpophalangeal joints; MRI: magnetic resonance imaging; RA: rheumatoid arthritis; PsA: psoriatic arthritis; PPV: positive predictive value; NPV: negative predictive value; TJC: tender joint count.

palpation for joint swelling and synovitis or edema MRI scores (Table 2).

DISCUSSION

This is the first study to correlate clinical examination findings with MRI of individual joints in both RA and PsA. Our study makes the following clinically relevant and important observations. First, a wide degree of variability among experienced examiners exists in the clinical assessment of inflammatory arthritis, particularly as related to joint swelling. Second, joint tenderness and stress pain in RA and PsA are more strongly associated with MRI synovitis and edema scores than is joint swelling. Finally, palpation and visual inspection of MCP joints of patients with PsA

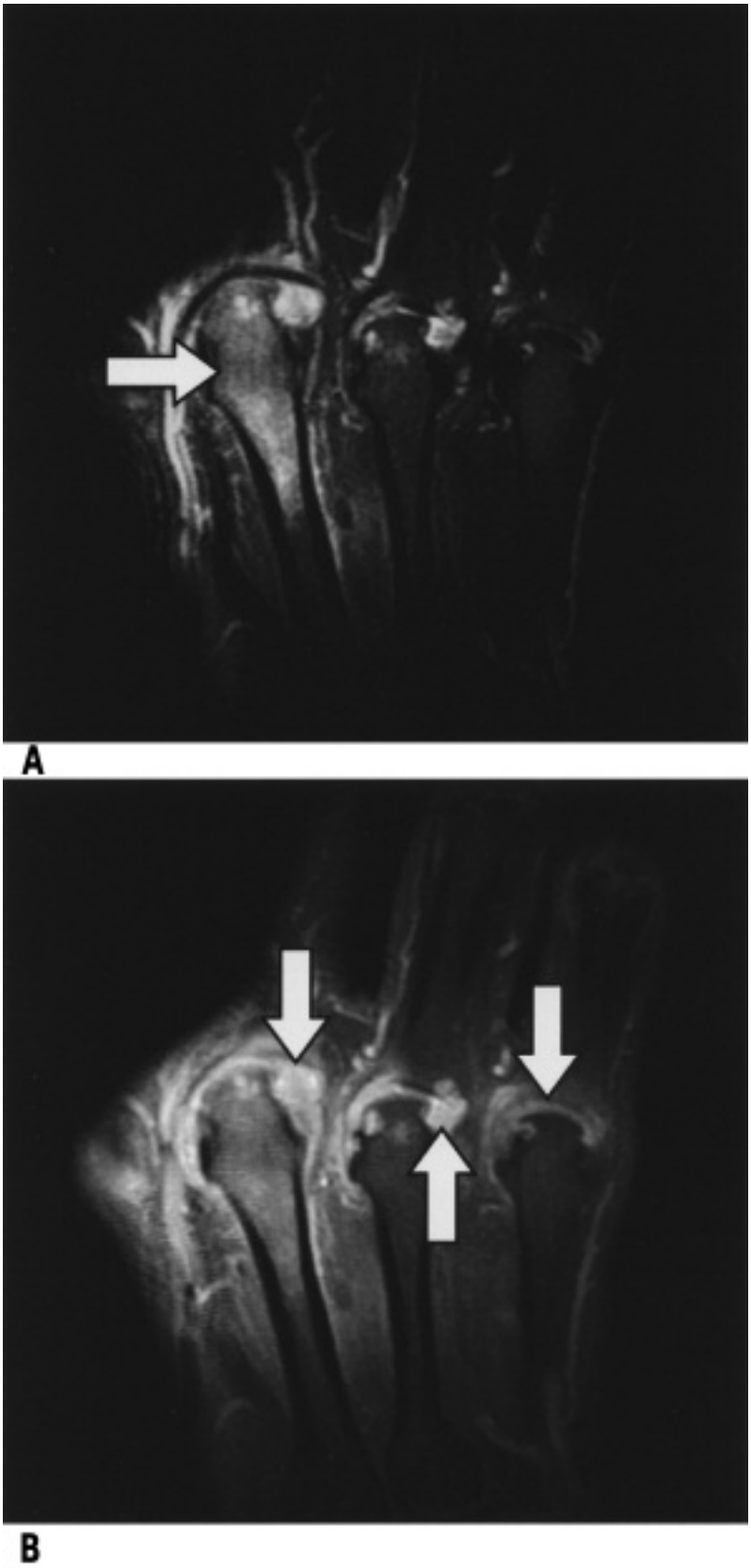

Figure 2. T2-weighted fat-suppressed (A) and post-contrast T1-weighted fat-suppressed (B) MR images illustrating the 3rd through 5th metacarpophalangeal (MCP) joints. Coronal T2-weighted image (A) illustrates associated bone edema most prominently seen within the 3rd metacarpal (white arrow). Coronal post-contrast T1-weighted image with fat suppression (B) shows marked intraarticular enhancement (arrows) consistent with synovitis of the 3rd through 5th MCP joints.

were unreliable and not sensitive in detecting synovial pathology.

In our study interrater reliability between examiners of palpation for joint swelling and tenderness in RA was modest at best, while visual inspection of swelling was considerably more reliable. The lack of interrater reliability of joint swelling in inflammation in PsA has been demonstrated by Gladman, *et al*[12,13]. We demonstrated a good agreement for the TJC but a moderate agreement for swollen joints. The poor interrater reliability of palpable joint swelling may have accounted for the rather modest ability of joint swelling to assess synovitis (as determined by MRI).

In RA, surprisingly, the most reliable method of determining synovial pathology by MRI was visual assessment of MCP joint swelling, which also had a PPV for the detection of synovitis in the joint. This was not unexpected, since visual findings of synovitis suggest more substantial synovitis compared to joints requiring palpation to detect synovitis. There was also a good association of joint-line tenderness with synovitis and erosion scores in both RA and PsA. Our finding of an association between tenderness and synovitis on MRI for RA and tenderness with synovitis and edema on MRI in PsA is supported by previous studies[4,14]. For example, Bond, *et al* demonstrated that tender joints were also predictive of radiological progression[4]. Our results suggest that visual inspection of swelling but not palpation may be a better guide for therapeutic decision-making in RA, and that tender joint assessment was superior to the assessment of swelling given variability between physicians in detecting palpable swelling. These results confirm previous data of better interrater reliability of tender rather than swollen joints[8]. However, a potential weakness of our study is that the assessors, although experienced in joint examination, did not undergo formal training prior to the assessments. Klinkhoff, *et al* have shown that training actually improves agreement among rheumatologists[15]. This is a factor to be considered in the design of larger studies that aim to validate our findings.

We demonstrated a lack of reliability to predict synovitis by assessment of joint swelling in PsA. This may reflect a predominance of periarticular capsular thickening rather than intraarticular synovial proliferation. Thus "joint thickening" without ballottable swelling might represent the extraarticular pathology. In contrast, tender joints were associated with synovitis and joints that were not tender did not demonstrate synovitis on MRI. The evaluation of stress pain with edema on MRI gives further support to the concept that tender joint examination is more accurate at detecting disease activity in patients with PsA. The determination of tenderness alone in PsA would also greatly simplify the clinical examination, but this needs to be validated in larger studies before any definitive recommendations are made. Alternative modalities such as MRI and ultrasound may be important adjuncts to clinical evaluation in patients with PsA where uncertainty exists regarding clinical disease activity. Further studies of the utility of these imaging modalities are warranted.

In contrast to RA, where MRI findings correlate with miniarthoscopic histological findings[16], no such correlation has been performed in PsA. In this context, a potential weakness of this study is the use of MRI as a "gold standard for PsA." The relatively small sample size is also a drawback to accurate interpretation of the results.

The significant correlation between stress pain and bone edema in both RA and PsA should also be noted. This suggests that stress pain may represent the finding of active inflammation within the bone marrow and may be useful as a clinical indicator in this regard. This finding, however, contrasts with the low correlation between tender joints and radiographic progression.

Our study suggests that for less obvious joint swelling, palpation may not be a reliable measure of synovitis. As well, joint examination in PsA does not provide an accurate reflection of synovitis compared to RA. Further studies of sensitive imaging techniques are required in PsA.

REFERENCES

1. Boers M. Demonstration of response in rheumatoid arthritis patients who are nonresponders according to the American College of Rheumatology 20% criteria: the paradox of beneficial treatment effects in nonresponders in the ATTRACT trial. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy. Arthritis Rheum 2001;44:2703-4.
2. Boers M, Kostense PJ, Verhoeven AC, van der Linden S. Inflammation and damage in an individual joint predict further damage in that joint in patients with early rheumatoid arthritis. Arthritis Rheum 2001;44:2242-6.
3. Gerber LH, Furst G, Yarboro C, El-Gabalawy H. Number of active joints, not diagnosis, is the primary determinant of function and performance in early synovitis. Clin Exp Rheumatol 2003;21 Suppl:S65-70.
4. Bond SJ, Farewell VT, Schentag CT, Gladman DD. Predictors for radiological damage in psoriatic arthritis: Results from a single centre. Ann Rheum Dis 2007;66:370-6.
5. Landewe RB, Geusens P, van der Heijde DM, Boers M, van der Linden SJ, Garnero P. Arthritis instantaneously causes collagen type I and type II degradation in patients with early rheumatoid arthritis: a longitudinal analysis. Ann Rheum Dis 2006;65:40-4.
6. Buskila D, Langevitz P, Gladman DD, Urowitz S, Smythe HA. Patients with rheumatoid arthritis are more tender than those with psoriatic arthritis. J Rheumatol 1992;19:1115-9.
7. McGonagle D, Conaghan PG, O'Connor P, Gibbon W, Green M, Wakefield R et al. The relationship between synovitis and bone changes in early untreated rheumatoid arthritis: a controlled magnetic resonance imaging study. Arthritis Rheum 1999; 42:1706-11.
8. Thompson PW, Hart LE, Goldsmith CH, Spector TD, Bell MJ, Ramsden MF. Comparison of four articular indices for use in clinical trials in rheumatoid arthritis: patient, order and observer variation. J Rheumatol 1991;18:661-5.
9. 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.

10. Lassere M, McQueen F, Ostergaard M, Conaghan P, Shnier R, Peterfy C, 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.
11. Kuritz SJ, Landis JR, Koch GG. A general overview of Mantel-Haenszel methods: applications and recent developments. Annu Rev Public Health 1988;9:123-60.
12. Gladman DD, Cook RJ, Schentag C, Feletar M, Inman RD Hitchon C, et al. The clinical assessment of patients with psoriatic arthritis: results of a reliability study of the Spondyloarthritis Research Consortium of Canada. J Rheumatol 2004;31:1126-31.
13. Gladman DD, Inman RD, Cook RJ, Maksymowych WP, Braun J, Davis J, et al. INternational SPondyloarthritis Inter-Observer Reliability Exercise — The INSPIRE Study: II. Assessment of peripheral joints, enthesitis and dactylitis. J Rheumatol 2007;34:1740-5.
14. Szkudlarek M, Narvestad E, Klarlund M, Court-Payen, Thomsen HS, Ostergaard M. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography, and clinical examination. Arthritis Rheum 2004;50:2103-12.
15. Klinkhoff AV, Bellamy N, Bombardier C, Carette S, Chalmers A, Esdaile JM, et al. An experiment in reducing interobserver variability of the examination for joint tenderness. J Rheumatol 1988;15:492-4.
16. Ostendorf B, Peters R, Dann P, Becker A, Scherer A, Wedekind F, et al. Magnetic resonance imaging and miniarthroscopy of metacarpophalangeal joints: sensitive detection of morphologic changes in rheumatoid arthritis. Arthritis Rheum 2001;44:2492-502.