Operational Definitions and Observer Reliability of the Plain Radiographic Features of Psoriatic Arthritis

WILLIAM J. TAYLOR, GUY G. PORTER, and PHILIP S. HELLIWELL

ABSTRACT. ***Objective.*** To determine the standardization and the observer reliability of potentially diagnostic plain radiographic features of psoriatic arthritis (PsA).

Methods. Radiographic features were chosen on the basis of a systematic review of the literature. One hundred sixty-four radiographs from 62 patients were selected from various sources by a musculoskeletal radiologist. Radiographs were read independently by 2 observers (WJT, PSH) and scored for the presence or absence of each evaluated feature. Cohen's kappa was used to determine observer agreement beyond chance, and the accuracy of each observer (with reference to the radiologist's judgment) was determined by likelihood ratios.

Results. The 2 observers demonstrated similar accuracy, although WJT tended to be more accurate for items classed as absent and PSH more accurate for items classed as present. The following features showed sufficient reliability to be reasonably included in further testing of their discriminatory value (intra- and interobserver kappa values): marginal syndesmophyte (0.68, 0.69), non-marginal syndesmophyte (0.75, 0.59), paravertebral ossification (0.89, 0.79), destructive discovertebral lesion (0.85, 0.65), Romanus lesion (0.64, 0.43), sacroiliitis (0.99, 0.86), entheseal erosion (0.80, 0.71), entheseal ossification (0.69, 0.76), distal interphalangeal erosive disease (0.58, 0.52), joint osteolysis (0.62, 0.47), juxtaarticular bony proliferation (0.43, 0.42), bony ankylosis (0.53, 0.54), tuft osteolysis (0.51, 0.36). The features that showed inadequate reliability were: loss of cortical definition of terminal tuft (0.33, 0.31) and periosteal new bone formation (0.42, 0.03).

Conclusion. A number of plain radiographic features of PsA have sufficient reliability to justify inclusion in diagnostic classification criteria sets for further testing. (J Rheumatol 2003;30:2645–58)

Key Indexing Terms:
PSORIATIC ARTHRITIS RADIOGRAPHY OBSERVER VARIATION

Despite several proposed classification criteria, any current case definition of psoriatic arthritis (PsA) has yet to be properly validated[1]. This poses substantial problems to the clinical investigator and renders the interpretation of clinical studies of PsA very difficult[2]. Several proposed classification criteria include radiographic features[3-6]. Although not all authors agree[7], there is evidence that certain axial radiological features are more common in PsA than in ankylosing spondylitis (AS)[8-10], and certain peripheral features are more common than in rheumatoid arthritis (RA)[4,11]. There is also increasingly strong evidence for the dominant role of enthesitis or osteitis in the pathology of spondyloarthropathies (SpA) to the extent that radiological evidence of enthesitis has been suggested as a discriminatory feature from RA[12]. However, it remains unclear exactly how discriminative any radiographic feature may be, particularly given their relative infrequency[13,14].

A further problem is that the literature concerning the radiology of PsA is somewhat confusing in its terminology. For example, there are several ways of describing osteolytic changes, including "resorptive arthritis[15]," "mutilation[16]," "mushrooming," "cup in stem[13]," "pencil in cup deformity[14]," "whittling of terminal phalanges," "pseudowidening of the interosseous joint space[17]," "phalangeal tuft resorption," or "osteolysis producing a widely, sharply demarcated joint space[11]." Similarly, there is some potential confusion regarding the correct meaning of "non-marginal syndesmophytes" with seeming equivalence between "paramarginal[18]," "parasyndesmophyte[17]," "non-marginal[10]," "comma shaped[10]," and "chunky[8]." If plain radiographic features are to be used in classification criteria, then clear definitions and standardization of terminology is important.

A third potential problem with using radiological imaging as part of classification criteria is the observer reliability of such features. While there is some literature concerning the inter- and intrarater reliability for sacroiliitis grading in AS[19,20], there are no reported data concerning the observer reliability of other radiological features of PsA.

We report the results of a systematic review of the litera-

From the Rheumatology and Rehabilitation Research Unit, University of Leeds, Leeds, United Kingdom; and the Department of Radiology, Airedale General Hospital, Keighley, West Yorkshire, United Kingdom.

Dr. Taylor was the recipient of the Dorothy Eden Fellowship from the Arthritis Foundation of New Zealand and the Arthritis Research Campaign (Grant No. T0520).

W.J. Taylor, MBChB, FRACP, FAFRM, Research Fellow; P.S. Helliwell, MD, PhD, FRCP, Senior Lecturer, Rheumatology and Rehabilitation Research Unit, University of Leeds; G.G. Porter, MD, FRCR, Consultant Radiologist, Department of Radiology, Airedale General Hospital.

Address reprint requests to Dr. W.J. Taylor, Rehabilitation Teaching and Research Unit, Department of Medicine, Wellington School of Medicine, University of Otago, PO Box 7343, Wellington, New Zealand.
E-mail: wtaylor@wnmeds.ac.nz

Submitted April 8, 2002; revision accepted May 13, 2003.

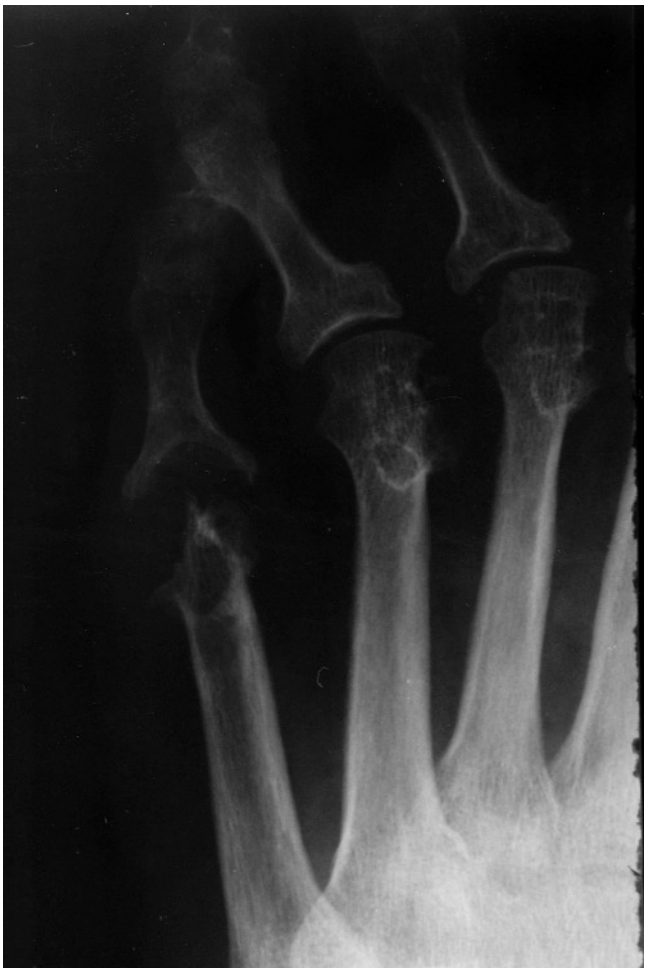

Figure 1. View of 5th toe metatarsophalangeal joint showing the classic "pencil in cup" deformity that we have termed "joint osteolysis."

ture in order to conceptualize and standardize operational definitions for the plain radiological features of PsA, and the results of an observer reliability study for the presence or absence of these features. We emphasize that the objective of this study was to determine whether standardized radiographic features could be reliably observed; whether such radiographic features truly distinguish PsA is the focus of a subsequent study. This work is part of preparation towards a multicenter prospective validation of classification criteria for PsA.

MATERIALS AND METHODS

We searched the relevant chapters of standard rheumatology and radiology textbooks and the source references were supplemented by personal archives and a Medline search (1966 to September 2000) using the search terms "psoriatic arthritis" and "radiography" as MeSH terms and as textwords. Eligibility for article selection and examination were: English language, PsA was the primary disorder under investigation or discussion, and the appearances of plain radiographs were an important focus. The literature was summarized by one author (WJT) and following discussions and a review meeting of all 3 authors, a standard system of terminology and definitions was agreed upon. A training meeting confirmed the radiographic appearances of these features.

Using these definitions, one author (GGP, a musculoskeletal radiologist) selected examples of each radiological feature from a number of sources including personal archives, clinical colleagues, and SpA clinics. The patients did not necessarily have PsA, since it was the selected radiological feature and not the diagnostic accuracy that was being evaluated in this study. More than one radiograph could be evaluated from a single patient and each radiograph could reveal more than one feature. Each radiograph was independently evaluated for presence or absence of applicable radiological features by 2 rheumatologists (WJT, PSH). Each observer has at least 5 years of clinical rheumatology experience and one (PSH) has previously published research regarding the plain radiology of PsA. One observer (WJT) repeated the evaluation of radiographs after a period of 2 weeks to calculate intraobserver reliability.

Concordance was evaluated in terms of the presence or absence of each applicable feature on the radiograph as a whole and not for an individual joint. This means that it was possible for agreement to have occurred if a feature was deemed to be present by one observer at one joint, but deemed to be present by the second observer at another joint (both joints visible on the same radiograph). Joint-by-joint data were not collected. Since the diagnostic context does not require the presence of the feature at a specific joint, but within groups of joints (e.g., small joints of hands and feet), we felt that this approach was valid.

Conventionally Cohen's kappa statistic is used to examine the chance-corrected agreement. We aimed to approximately limit the 95% confidence interval of kappa such that the lower bound was not less than 0.4 when observed agreement was 85% and prevalence of the feature was 30% to 60%, which suggested an approximate sample size of at least 30 radiographs. We added 0.1 to cells with zero values in order to calculate all indices. Since bias (differential probability of each observer to report a positive finding) and prevalence (overall frequency of the radiological feature) affect kappa, so as to obscure the measurement of chance-corrected agreement, we report the bias index, prevalence index, and adjusted kappa suggested by Bryt[21]. The bias index reflects the extent to which there is a difference between observers in the probability of a positive rating when the observer is uncertain. Cohen's kappa assumes this to be equal. When this is not so, kappa may not reflect true chance-corrected agreement. The prevalence index reflects the effect that very high or very low frequencies of the feature can have on distorting kappa as an unbiased measure of chance-corrected agreement. This is analogous to the effect that very high or very low prevalence of disease has on the positive predictive value of diagnostic tests, despite such tests having high accuracy (specificity and sensitivity). In this context prevalence means the frequency of the feature among evaluable radiographs, and not the frequency of the feature among people with PsA. Following the suggestions by Cicchetti and Feinstein[22], we also report the proportionate agreement in the observers' positive and negative decisions (the number of instances of both observers agreeing on the presence/absence of a feature divided by the average number of positive/negative readings by both observers). These indices help illuminate reasons for inappropriately low or high kappa values, since there may be better agreement about the absence of a feature than the presence of a feature, or vice versa. SPSS (v 7) and DAG_STAT[23] were used for the analysis.

RESULTS

Systematic literature review. A total of 122 articles were identified by the Medline search, of which 25 fulfilled the predefined criteria and were examined. A further 42 articles were identified from references and textbook bibliographies. Six standard rheumatology and radiology textbook chapters were also reviewed. A total of 73 articles were examined, including 37 uncontrolled case-series or reports, 20 controlled case-series, and 16 review articles (including

Table 1. Operational definitions of the principal plain radiograph features of psoriatic arthritis.

Feature	Definition	Key References
Axial Features		
Marginal syndesmophyte	Classic, thin syndesmophyte arising vertically from annular attachment to vertebral body	8, 9, 33
Non-marginal syndesmophyte	Vertically oriented or curvilinear syndesmophyte, often thick and chunky, arising from beyond the annular attachment to vertebral body	8–10, 33
Paravertebral ossification	Ossification close to vertebral body, but with a clearly defined gap between the margins of the ossification and the vertebral body	8, 10, 31–33
Destructive discovertebral lesion	Irregularity of superior and inferior endplates with erosive changes and adjacent vertebral sclerosis, with or without fracture and angulation (Andersson lesion)	34, 35, 44
Romanus lesion	Clearly defined erosion of the anterior margin of the discovertebral junction at the superior or inferior portions of the vertebral body	45, 46
Sacroiliitis	New York grade $\geq$ 2 bilaterally, or $\geq$ 3 unilaterally	39, 47, 48
Peripheral Features		
Extraarticular entheseal erosion	Erosion at entheseal insertion of calcaneus, ischial tuberosities, iliac crest, femoral trochanters, humeral tuberosity, or patella	49, 50
Extraarticular entheseal ossification	Irregular bony proliferation at entheseal insertion of calcaneus, ischial tuberosities, iliac crest, femoral trochanters, humeral tuberosity, patella	49–51
DIP erosive disease, excluding erosive OA	Clearly defined marginal erosion of DIP joint AND either: evidence of joint destruction (widened joint space or osteolysis) or juxtaarticular periostitis OR absence of osteophytes, joint space narrowing or central erosive change	11, 13, 14, 52
Joint osteolysis	Osteolysis producing a wide, sharply demarcated joint space, including phalangeal whittling	4, 11, 13
Tuft osteolysis	Osteolysis of terminal phalangeal tuft	4, 11–14, 48, 49
Loss of tuft cortical definition	Loss of cortical definition of terminal tuft, often with a "fluffy" appearance	4
Juxtaarticular bony proliferation	Ill-defined ossification near joint margins, but excluding osteophyte formation	4, 9, 51, 53, 54
Periosteal new bone formation	Linear, ill-defined metaphyseal or diaphyseal bony apposition	9, 29, 49, 51, 55
Bony ankylosis	Bony ankylosis indicated by trabeculae crossing the joint space	4, 11, 13, 49

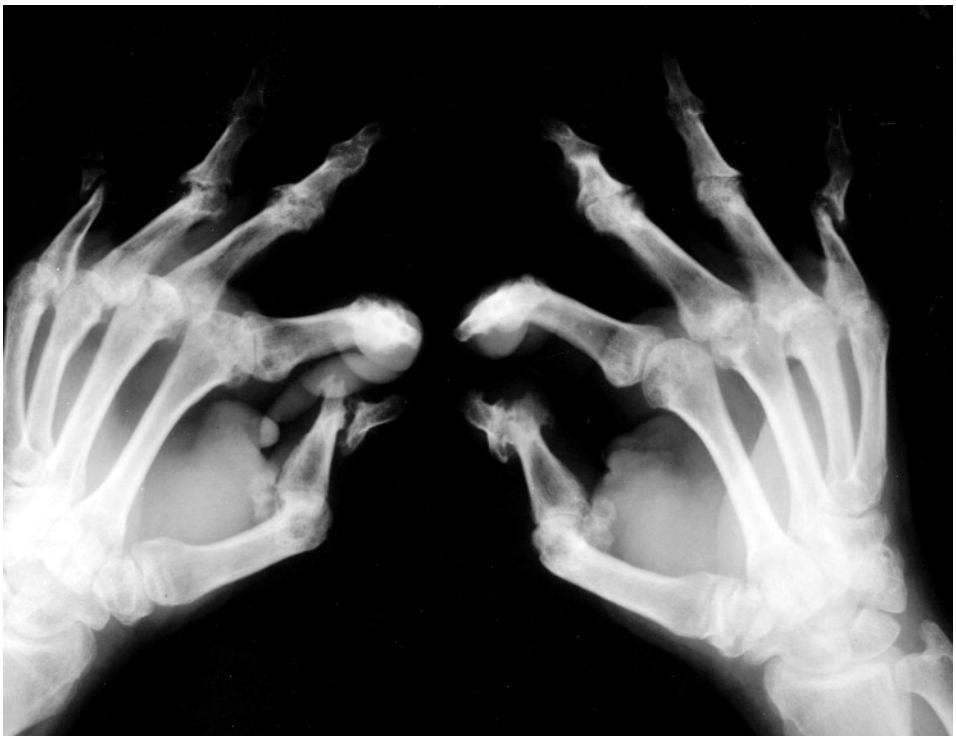

Figure 2. Ball-catcher's view of both hands shows DIP erosive disease, loss of tuft cortical definition, juxtaarticular bony proliferation, bony ankylosis, periosteal new bone formation, and joint osteolysis.

Table 2. Agreement between the 2 observers (WJT, PSH) and the radiologist's determination of the presence/absence of each radiological feature.

Feature (frequency of feature according to the radiologist rating)		Observed Agreement with Radiologist (95% CI)	Kappa (95% CI)
Axial Features			
Marginal syndesmophyte (0.30)	WJT	0.88 (0.78, 0.95)	0.73 (0.55, 0.90)
	PSH	0.83 (0.72, 0.91)	0.59 (0.38, 0.80)
Non-marginal syndesmophyte (0.31)	WJT	0.78 (0.66, 0.87)	0.47 (0.24, 0.70)
	PSH	0.82 (0.70, 0.90)	0.56 (0.34, 0.78)
Destructive discovertebral lesion (0.10)	WJT	0.96 (0.88, 1.00)	0.78 (0.49, 1.00)
	PSH	0.93 (0.83, 0.98)	0.47 (0.05, 0.90)
Romanus lesion (0.15)	WJT	0.81 (0.69, 0.91)	0.27 (–0.07, 0.60)
	PSH	0.89 (0.77, 0.96)	0.56 (0.25, 0.87)
Paravertebral ossification (0.06)	WJT	0.92 (0.83, 0.97)	0.25 (0.00, 0.70)
	PSH	0.94 (0.85, 0.98)	0.30 (0.00, 0.80)
Sacroiliitis (0.67)	WJT	0.96 (0.85, 0.99)	0.90 (0.77, 1.00)
	PSH	0.90 (0.78, 0.97)	0.79 (0.61, 0.96)
Peripheral Features			
Entheseal erosion (0.34)	WJT	0.71 (0.53, 0.85)	0.26 (0.00, 0.60)
	PSH	0.81 (0.64, 0.93)	0.55 (0.24, 0.85)
Entheseal ossification (0.53)	WJT	0.78 (0.61, 0.90)	0.57 (0.29, 0.83)
	PSH	0.76 (0.59, 0.89)	0.53 (0.24, 0.81)
DIP erosive disease (0.48)	WJT	0.69 (0.54, 0.80)	0.37 (0.12, 0.62)
	PSH	0.67 (0.52, 0.79)	0.32 (0.09, 0.56)
Joint osteolysis (0.36)	WJT	0.75 (0.62, 0.86)	0.51 (0.28, 0.73)
	PSH	0.87 (0.75, 0.95)	0.69 (0.48, 0.89)
Tuft osteolysis (0.056)	WJT	0.91 (0.79, 0.97)	0.50 (0.15, 0.86)
	PSH	0.94 (0.85, 0.99)	0.37 (0.00, 0.93)
Loss of cortical definition of terminal tuft (0.50)	WJT	0.74 (0.60, 0.85)	0.48 (0.25, 0.71)
	PSH	0.63 (0.49, 0.76)	0.26 (0.04, 0.48)
Periosteal new bone (0.41)	WJT	0.56 (0.41, 0.69)	0.05 (0.00, 0.32)
	PSH	0.66 (0.52, 0.78)	0.20 (0.02, 0.37)
Juxtaarticular bony proliferation (0.63)	WJT	0.70 (0.56, 0.82)	0.39 (0.14, 0.64)
	PSH	0.67 (0.53, 0.79)	0.37 (0.14, 0.58)
Bony ankylosis (0.24)	WJT	0.82 (0.69, 0.91)	0.58 (0.36, 0.80)
	PSH	0.91 (0.78, 0.97)	0.75 (0.55, 0.96)

Table 3. Interobserver agreement between WJT and PSH (95% CI).

Feature (frequency of feature according to the radiologist rating)	Observed Agreement	Kappa	Positive Agreement	Negative Agreement
Axial Feature				
Marginal syndesmophyte (0.30)	0.87 (076, 0.94)	0.69 (0.50, 0.87)	0.78 (0.64, 0.92)	0.90 (0.84, 0.97)
Non-marginal syndesmophyte (0.31)	0.83 (0.72, 0.91)	0.59 (0.37, 0.80)	0.70 (0.54, 0.87)	0.88 (0.82, 0.95)
Destructive discovertebral lesion (0.10)	0.97 (0.89, 1.00)	0.65 (0.21, 1.00)	0.67 (0.23, 1.00)	0.98 (0.96, 1.00)
Romanus lesion (0.15)	0.88 (0.78, 0.95)	0.43 (0.10, 0.76)	0.50 (0.20, 0.80)	0.93 (0.88, 0.98)
Paravertebral ossification (0.06)	0.98 (0.92, 1.00)	0.79 (0.40, 1.00)	0.80 (0.42, 1.00)	0.99 (0.98, 1.00)
Sacroiliitis (0.67)	0.94 (0.83, 0.99)	0.86 (0.70, 1.00)	0.95 (0.90, 1.00)	0.90 (0.79, 1.00)
Peripheral Features				
Entheseal erosion (0.34)	0.91 (0.76, 0.98)	0.71 (0.41, 1.00)	0.77 (0.52, 1.00)	0.94 (0.88, 1.00)
Entheseal ossification (0.53)	0.88 (0.72, 0.97)	0.76 (0.54, 0.98)	0.88 (0.77, 1.00)	0.88 (0.75, 1.00)
DIP erosive involvement (0.48)	0.76 (0.62, 0.86)	0.52 (0.31, 0.73)	0.70 (0.54, 0.85)	0.80 (0.69, 0.91)
Joint osteolysis (0.36)	0.74 (0.60, 0.85)	0.47 (0.27, 0.67)	0.63 (0.45, 0.81)	0.80 (0.70, 0.90)
Tuft osteolysis (0.056)	0.89 (0.77, 0.96)	0.36 (0.00, 0.73)	0.40 (0.02, 0.78)	0.94 (0.89, 0.99)
Loss of cortical definition of terminal tuft (0.50)	0.63 (0.49, 0.76)	0.31 (0.12, 0.50)	0.55 (0.37, 0.72)	0.69 (0.56, 0.82)
Periosteal new bone (0.41)	0.65 (0.51, 0.77)	0.03 (0.00, 0.17)	0.10 (0.00, 0.27)	0.78 (0.69, 0.88)
Juxtaarticular bony proliferation (0.63)	0.70 (0.56, 0.82)	0.42 (0.19, 0.65)	0.69 (0.55, 0.84)	0.71 (0.58, 0.85)
Bony ankylosis (0.24)	0.80 (0.66, 0.89)	0.54 (0.32, 0.77)	0.69 (0.51, 0.86)	0.85 (0.76, 0.94)

Table 4. Interobserver agreement between WJT and PSH showing bias and prevalence-adjusted indices.

Feature (frequency of feature according to the radiologist rating)	Bias Index*	Prevalence Asymmetry Index**	Bias Adjusted Kappa	Prevalence and Bias Adjusted Kappa
Axial Features				
Marginal syndesmophyte (0.30)	–0.07	0.39	0.68	0.73
Non-marginal syndesmophyte (0.31)	–0.05	0.44	0.59	0.67
Andersson lesion (0.10)	–0.03	0.91	0.65	0.94
Romanus lesion (0.15)	0.00	0.76	0.43	0.76
Sacroiliitis (0.68)	–0.06	–0.35	0.86	0.88
Paravertebral ossification (0.06)	–0.02	0.92	0.79	0.97
Peripheral Features				
Entheseal erosion (0.34)	0.03	0.61	0.71	0.82
Entheseal ossification (0.53)	0.06	–0.03	0.76	0.76
DIP erosive disease (0.48)	–0.20	0.20	0.50	0.52
Joint osteolysis (0.36)	–0.26	0.30	0.43	0.48
Tuft osteolysis (0.056)	–0.11	0.81	0.34	0.77
Loss of cortical definition of terminal tuft (0.50)	–0.33	0.19	0.23	0.26
Periosteal new bone (0.41)	–0.31	0.61	–0.12	0.30
Juxtaarticular bony proliferation (0.63)	–0.15	0.04	0.41	0.41
Bony ankylosis (0.24)	–0.13	0.35	0.54	0.59

*Larger positive values indicate PSH more likely to rate a feature as present, larger negative values indicate WJT more likely to rate a feature as present.
** Larger values indicate that the low frequency of the feature will make kappa inappropriately low as an index of chance-corrected agreement.

Table 5. Intraobserver agreement (WJT), 95% CI.

Feature	Observed Agreement (%)	Kappa
Axial Features		
Marginal syndesmophyte	86.4 (75.7, 93.6)	0.68 (0.50, 0.87)
Non-marginal syndesmophyte	89.4 (79.4, 95.6)	0.75 (0.58, 0.92)
Paravertebral ossification	98.7 (92.2, 99.9)	0.89 (0.52, 1.15)
Destructive discovertebral lesion	98.5 (91.8, 99.9)	0.85 (0.56, 1.14)
Sacroiliitis	99.5 (98.3, 100.0)	0.99 (0.98, 1.00)
Romanus lesion	93.9 (85.2, 98.3)	0.64 (0.32, 0.96)
Peripheral Features		
Entheseal erosion	94.7 (82.3, 99.4)	0.80 (0.54, 1.06)
Entheseal ossification	84.2 (68.8, 94.0)	0.69 (0.48, 0.91)
DIP erosive disease	79.3 (65.9, 89.2)	0.58 (0.37, 0.80)
Joint osteolysis	81.0 (68.0, 90.6)	0.62 (0.41, 0.82)
Tuft osteolysis	88.5 (76.6, 95.6)	0.51 (0.17, 0.85)
Loss of cortical definition of terminal tuft	67.9 (53.7, 80.1)	0.33 (0.08, 0.58)
Juxtaarticular bony proliferation	71.7 (57.7, 83.2)	0.43 (0.18, 0.67)
Periosteal new bone	75.5 (61.7, 86.2)	0.42 (0.17, 0.67)
Bony ankylosis	77.4 (63.8, 87.7)	0.53 (0.29, 0.76)

textbook chapters). This led to operational definitions of 14 potentially discriminatory plain radiographic features of PsA (Table 1). We discuss particular issues regarding radiograph appearances and terminology more fully below. A complete bibliography is available from the authors.

Observer variation. In total 164 radiographs were examined from 62 patients. These included views of the lumbar and cervical spine (64 radiographs), pelvis (27), sacroiliac joints (9), hands/wrist (30), feet (24), and heels, knees and shoulders (4 radiographs each).

The accuracy of each observer with respect to the radiologist's "gold-standard" is shown in Table 2. Correct classification rates were similar between the 2 observers. However, PSH consistently achieved better accuracy for items classed as present and WJT achieved better accuracy for items classed as absent (data not shown). This is probably due to a bias effect whereby PSH rated items more conservatively (less likely to rate as positive) than WJT. This can also be seen in Tables 3 and 4, where the interobserver indices are shown. Agreement for items classed as

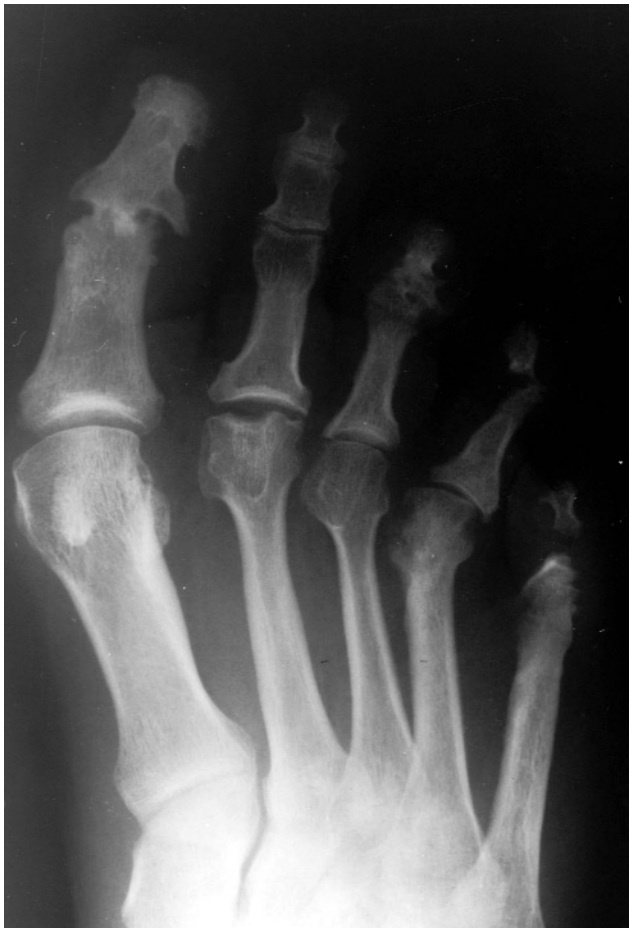

Figure 3. Posteroanterior view of the left foot shows DIP erosive disease, joint osteolysis, tuft osteolysis, loss of tuft cortical definition, and juxtaarticular bony proliferation.

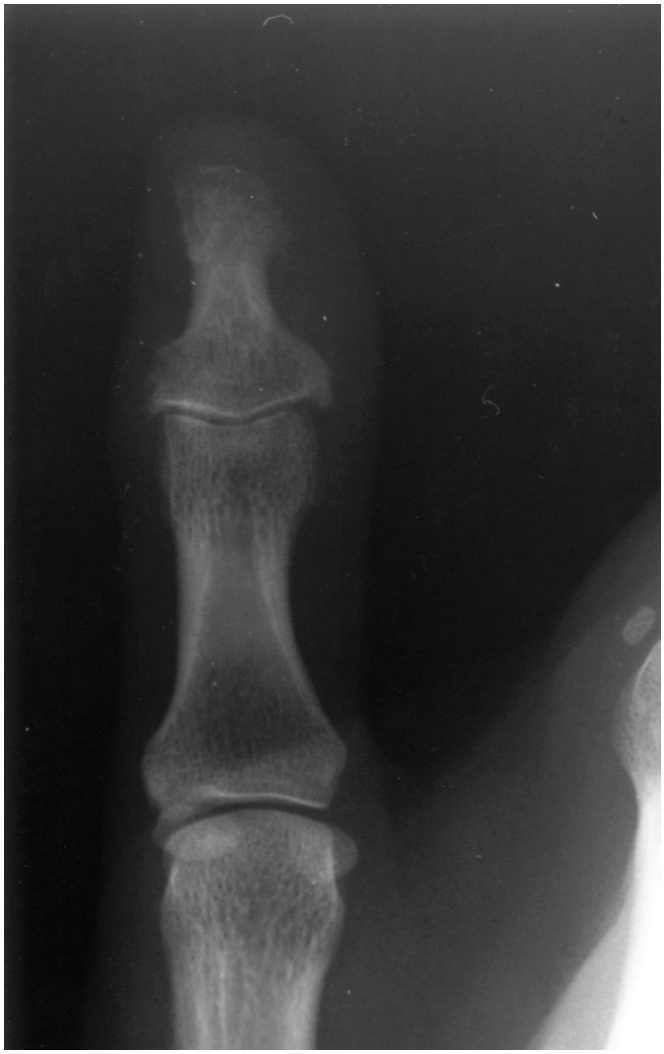

Figure 4. View of the left thumb shows juxtaarticular bony proliferation at the IP joint and loss of tuft cortical definition.

absent is generally greater than for items classed as present and there is at least a small bias index present for most items. Intraobserver reliability is shown in Table 5.

DISCUSSION

Selection and definitions. There were a number of instances in which osteolysis was reported to produce features characteristic of PsA. These include interphalangeal osteolysis producing a widened, sharply demarcated joint space[4], phalangeal tuft resorption[4], and whittling of bone ends[14]. Pencil in cup deformity is a combination of whittling on one side of the joint and extensive, deep erosion on the other side (Figure 1). There appeared to be no real difference between the process producing "pencil in cup" deformity and "interphalangeal osteolysis producing a widened, sharply demarcated joint space" and that the latter term was preferred since it more accurately describes joints for which "pencil in cup" seems inaccurate (Figures 2 and 3). For conciseness, we have named this feature "joint osteolysis." Although "pencil in cup" is a term strongly associated with PsA, the multiple terminology in the literature used to describe the same process made it sensible to choose a single term that encompassed these appearances. Since many instances of joint osteolysis occur that do not look like a "pencil in cup," especially in early stages, we suggest that "pencil in cup" is a special case of joint osteolysis and that the more general term is more suited for potential diagnostic criteria.

Since osteolysis of the terminal tuft resulting in "whittling" or erosion could be anatomically differentiated from interphalangeal joint osteolysis, we retained this as a specific anatomically distinct manifestation of osteolysis. We also attempted to distinguish between a clearly eroded terminal tuft (which we termed "tuft osteolysis"; Figure 3) and less clear-cut tuftal changes that we named "loss of tuft cortical definition" (Figure 4). Although the less severe changes have also been termed "tuftal osteolysis[24,25]," they have also been termed "osteoperiostitis of the distal

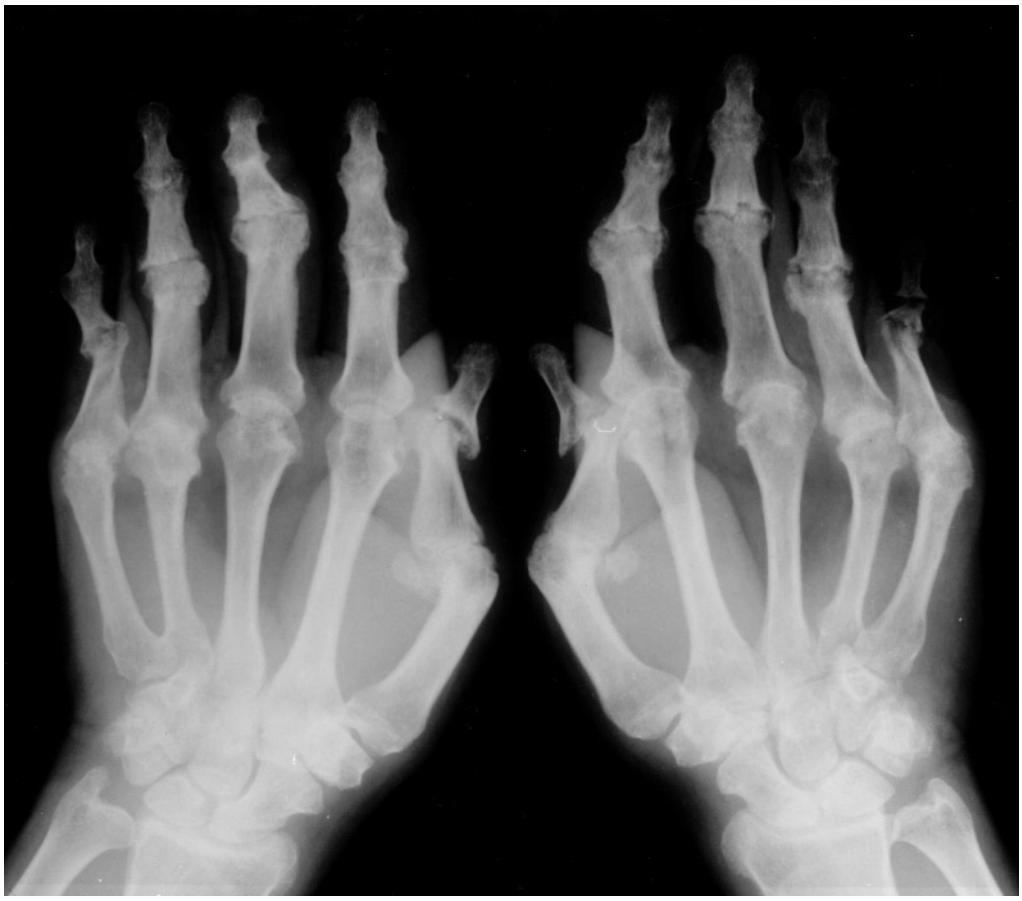

Figure 5. Posteroanterior view of same hands as in Figure 2, showing particularly the thickened phalanges due to periosteal new bone formation.

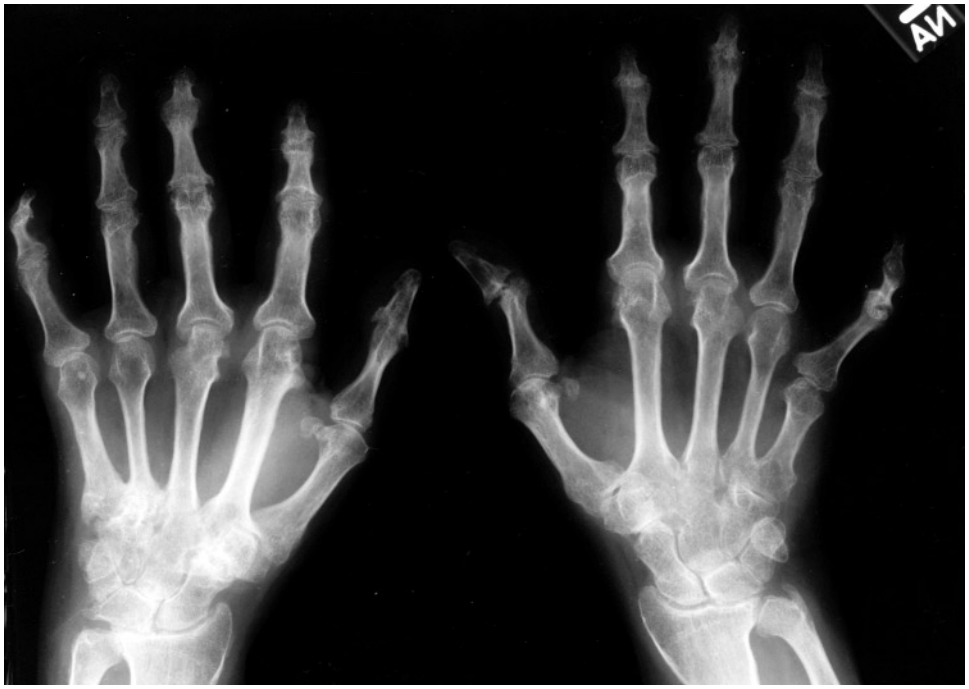

Figure 6. Both hands show features of DIP erosive disease, bony ankylosis, juxtaarticular bony proliferation, loss of tuft cortical definition, and periosteal new bone formation.

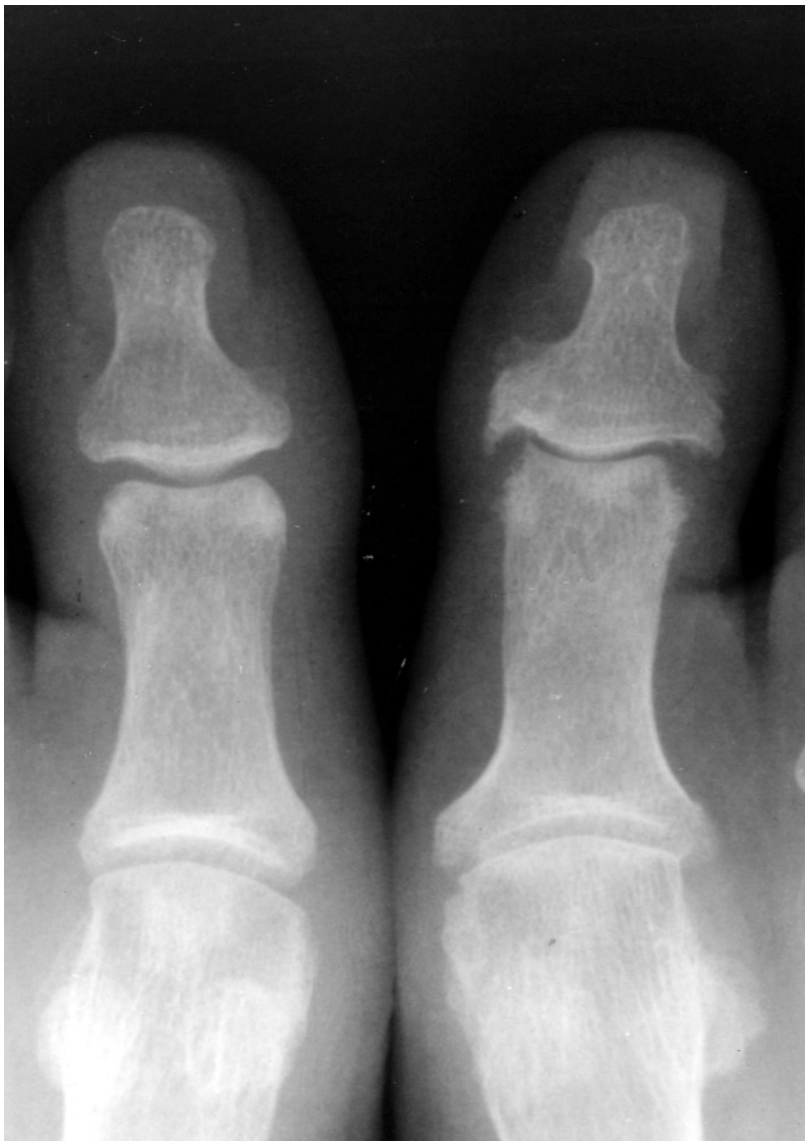

Figure 7. Posteroanterior view of both great toes showing marginal erosions and juxtaarticular bony proliferation at the left IP joint, to give the typical "mouse-ear" appearance.

phalanx[4]." As it is not entirely clear that these types of tuftal changes represent the same process (osteolysis or proliferation) and we felt that the less clear-cut changes may be subject to more observer variation, we regard the distinction as justifiable.

Marginal erosions at peripheral joints may occur in PsA, and since this may be distinguished from RA by preservation of juxtaarticular bone density, it might be possible to employ this characteristic in diagnostic criteria[3]. However, the assessment of juxtaarticular osteopenia is known to be highly dependent upon radiographic technique[26], so we chose only to evaluate erosions at distal interphalangeal (DIP) joints, in agreement with Fournie, *et al*[4] and Avila, *et al*[11].

Bony proliferation is a fairly characteristic feature of PsA, also occurring in other SpA, but rarely in RA[27]. Periostitis may be manifest as new bone formation close to and parallel to the cortex of phalanges, metacarpals, and metatarsals[28,29] (Figures 2, 5, 6). Other proliferative features include bony ankylosis, particularly of interphalangeal joints[11] (Figures 2 and 6), or juxtaarticular periostitis producing a spiculated or band-like image in a finger or toe[4] (Figures 4 and 7). Each of these sites is anatomically distinct and may be identified on plain radiographs. We have chosen to identify instances of new bone formation by anatomical site, rather than use terms such as "whiskering" or "paintbrush appearance." We did not evaluate the "ivory phalanx" appearance described by Resnick and Broderick[30], since we were unable to recognize sufficient examples of this feature in the radiographs available.

At axial sites, bony proliferation may be manifest as

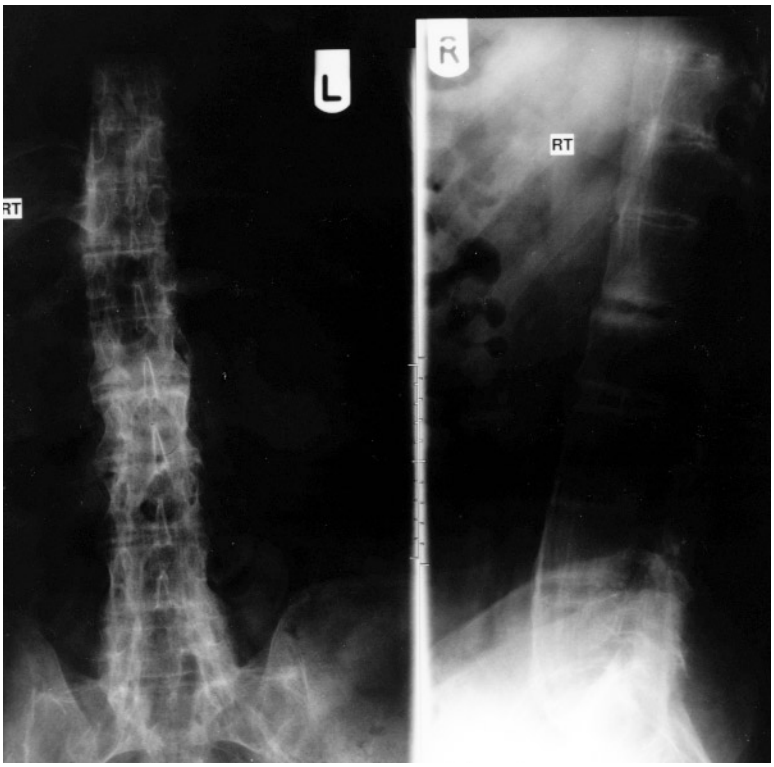

Figure 8. Anteroposterior and lateral spine showing marginal and non-marginal syndesmophytes and a destructive discovertebral lesion at L1/2.

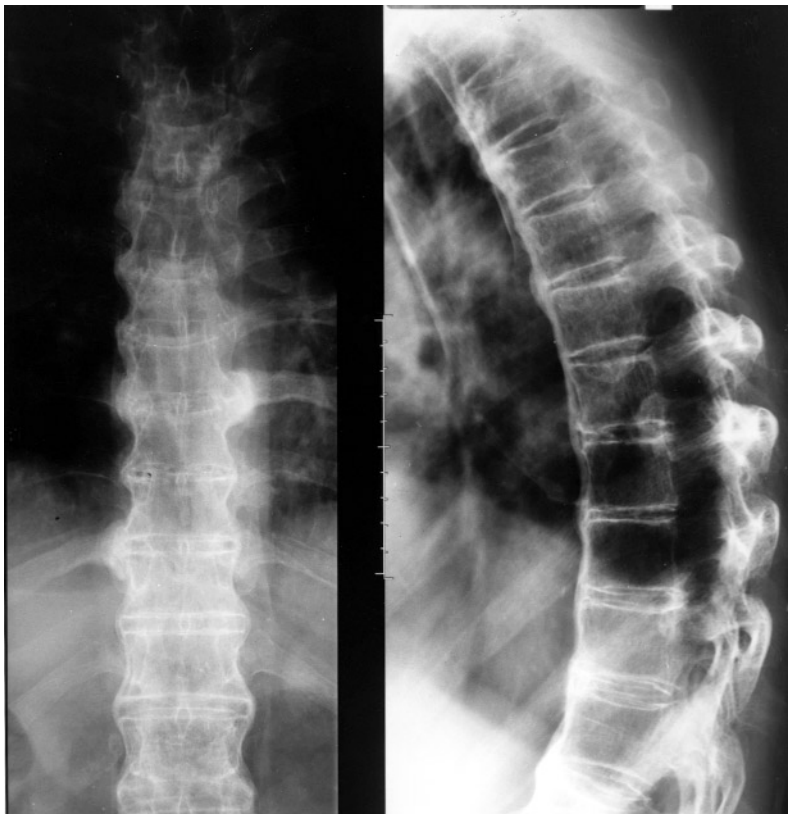

Figure 9. Anteroposterior and lateral thoracic spine showing all 3 morphologies of axial bony proliferation: marginal, non-marginal, and paravertebral ossification.

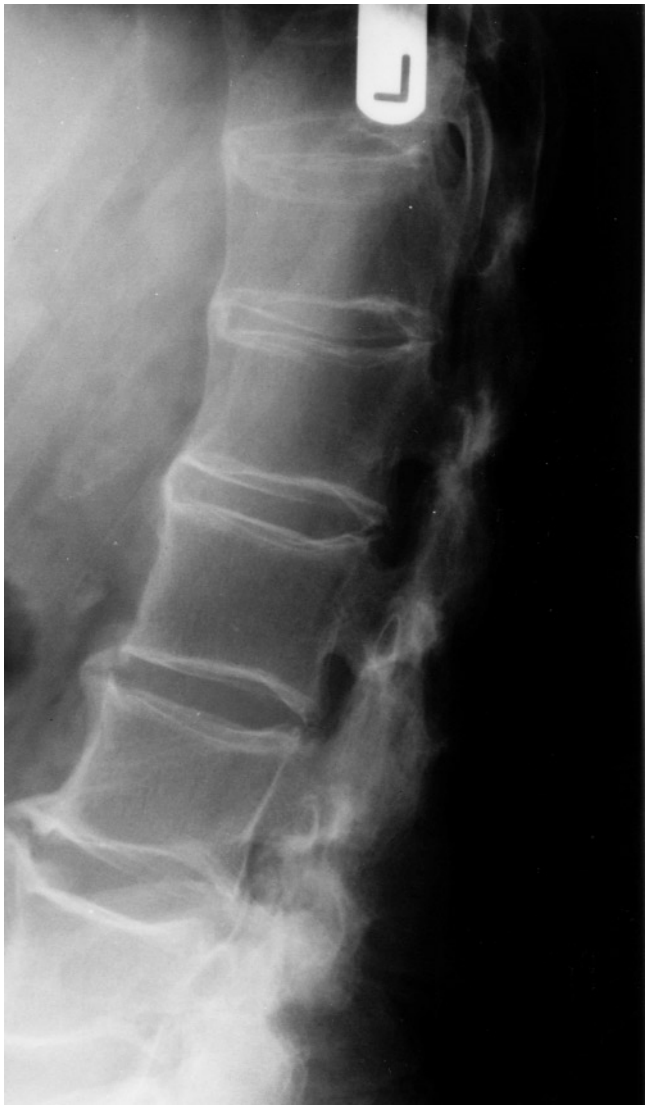

Figure 10. Lateral view of lumbar spine showing marginal and non-marginal syndesmophytes.

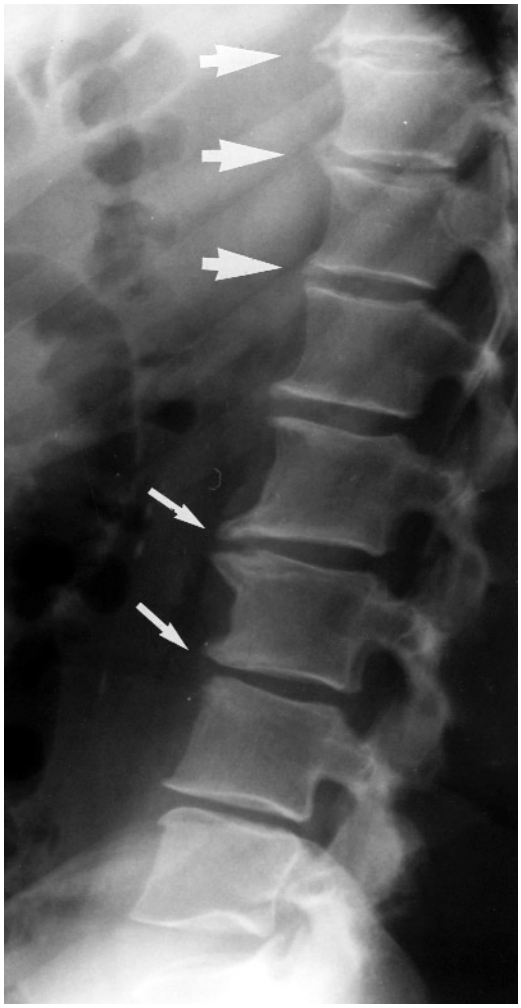

Figure 11. Lateral view of lumbar spine showing non-marginal syndesmophytes arising from beyond the annular attachment in a roughly vertical or curvilinear direction (large arrows), compared with osteophytes that are directed horizontally (small arrows).

paravertebral ossification, syndesmophytes, or sacroiliac ankylosis. There is some confusion in the literature regarding terminology: paravertebral ossification is said to be synonymous with paramarginal syndesmophytes[28], illustrated by a radiograph showing changes very similar to the "chunky syndesmophytes" of Helliwell, *et al*[8], which were defined as the "non-marginal" and "inverted comma syndesmophytes" described by McEwen, *et al*[10]. Conversely, Helliwell and Fournie define paravertebral ossification in relation to the report by Bywaters and Dixon[31], which describes ossification such that a distinct space exists between the linear areas of the ossification and the borders of the vertebral body. McEwen, *et al* refer to the various types of syndesmophytes as marginal, non-marginal, comma, and Bywaters-Dixon types[10]. It has been suggested that the Bywaters-Dixon lesion may progress to become a non-marginal syndesmophyte[32,33]. We conclude that syndesmophytes are of 3 morphologies: marginal, non-marginal (which includes "chunky" and "comma"), and paravertebral ossification (which we recognize may simply be an early stage of non-marginal). See Figures 8 to 11 for examples of these.

Discovertebral lesions such as osteitis, Romanus lesion, vertebral squaring, and the 5 types of lesion described by Cawley, *et al*[34], are said to be less common in psoriatic spondyloarthropathy than in AS[24]. To pursue this notion, we chose 2 clearly defined features — the Romanus lesion (Figure 12) and the destructive discovertebral lesion described by Andersson[35] (also classified as Cawley type *e*, or type III) — as representing discovertebral manifestations most likely to demonstrate satisfactory observer variation (Figure 8).

Although there are a number of other lesions of the

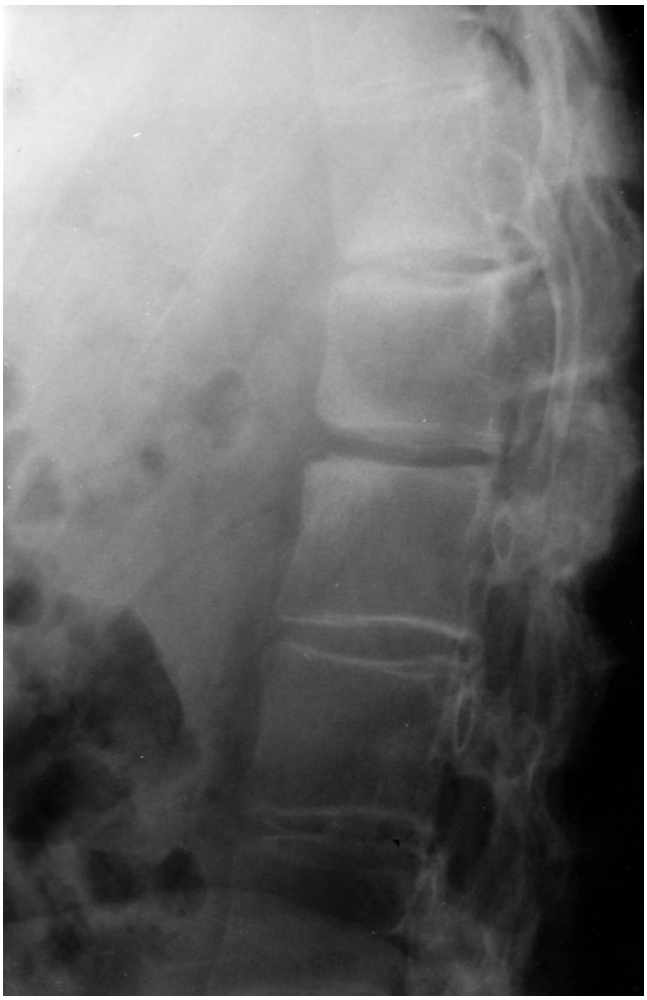

Figure 12. Lateral lumbar spine showing the Romanus lesion, a well defined erosion of the vertebral body at its annular attachment.

cervical spine reported including apophyseal joint and disc space narrowing, posterior ligamentous calcification, atlantoaxial subluxation, odontoid erosion, and subaxial erosions[15,36,37], we did not consider these to be sufficiently common or specific to proceed with further evaluation. None of these features have been suggested in published classification criteria.

Since sacroiliitis is not always manifest in PsA shown to involve the axial skeleton by the presence of syndesmophytes[32,39], it is important that sacroiliitis be evaluated as a separate criterion. Radiographic sacroiliitis has been defined using the New York grading system[39] and the radiological portion of the modified classification criteria for AS defines presence or absence of radiographic sacroiliitis: grade 2 to 4 bilaterally or grade 3 to 4 unilaterally[40].

Since enthesitis may be fundamentally important in the classification of inflammatory arthritis[41], magnetic resonance imaging (MRI) evidence of enthesitis has been suggested as an important discriminating feature for the diagnosis of PsA[5]. The discriminatory value of MRI may not be so good in well established disease, so that plain radiographic signs of enthesitis have been suggested as potentially more useful in discriminating RA-like PsA from RA (McGonagle D, personal communication). The radiographic features of enthesitis are erosion and/or ossification at entheseal insertions, which might include the following sites[42]: posterior and plantar aspect of the calcaneus, femoral trochanters, ischial tuberosities, ankle malleoli, distal portion of femoral condyles, olecranon of the ulna, iliac crest, inferior margin of clavicle, anterior portion of patella, and spinous processes of verterbrae. We pragmatically limited the observation sites to the pelvis, knees, heels, and shoulders (see Figures 13 and 14 for examples).

Observer variation. The conventional interpretation of Cohen's kappa is that values > 0.4 indicate at least "moderate" agreement[43]. Using this level of agreement as a benchmark, the following plain radiographic features exhibit too much observer variation to be useful as potential discriminatory items in classification criteria: loss of the cortical definition of the terminal phalangeal tuft and periosteal new bone formation. Some features were not sufficiently prevalent to be entirely confident about observer agreement (tuft osteolysis, paravertebral ossification, Romanus lesion, and destructive discovertebral lesion), although the prevalence-adjusted kappa index would suggest that there is at least substantial agreement for the presence of these features. Since terminal tuft changes may be highly specific for PsA, despite the infrequency and borderline satisfactory interobserver variation, it seems useful to retain this feature for further evaluation. There was "substantial" agreement (kappa > 0.6) for several features including marginal and non-marginal syndesmophytes, sacroiliitis, and entheseal changes, and "moderate" agreement (kappa > 0.4) for bony ankylosis, joint osteolysis, DIP erosive disease, and juxtaarticular bony proliferation. Observed agreement was at least 85% for most of these features.

It is unclear what level of agreement in this highly artificial scenario would be required to be confident about the reliability of these radiological features for case-definition in the field. Values for interobserver reliability of kappa 0.38 to 0.64 for the Bath Ankylosing Spondylitis Radiological Index and 0.37 to 0.47 for sacroiliac joint scores have been interpreted as being adequate for radiological scoring methods in AS[20]. To maintain satisfactory agreement, proper training in these methods is mandatory. To assist with standardization, an atlas of radiological appearances and standard viewing conditions may be useful.

The principal limitation to this study is the small number of highly selected radiographs. The natural prevalence of most of the radiological features that were examined is not known, but is certain to be less frequent than was observed

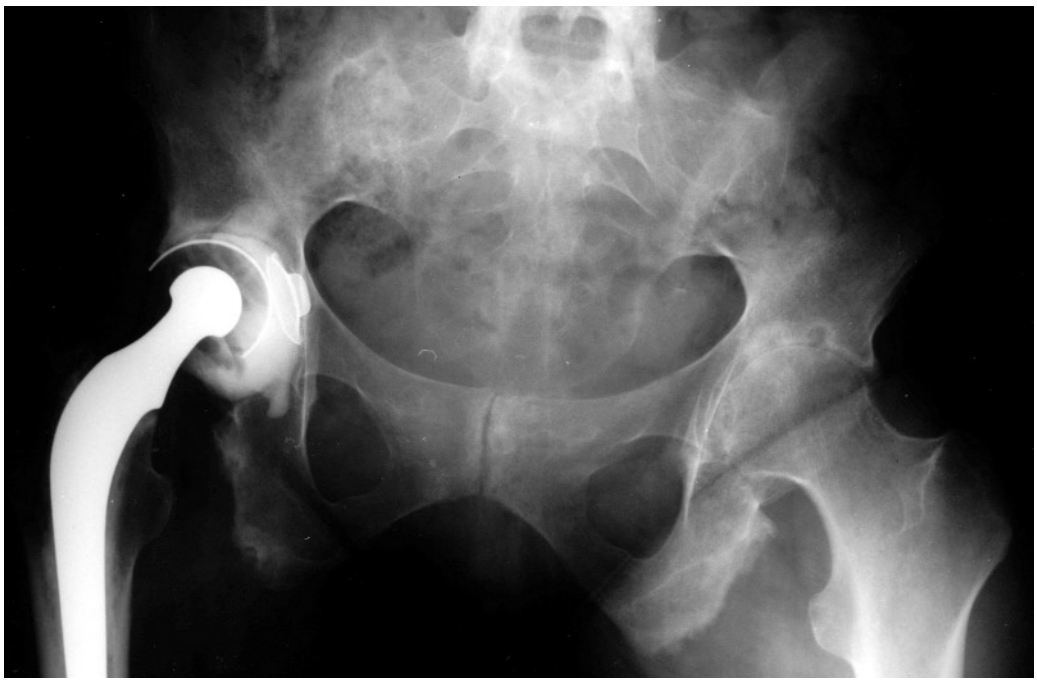

Figure 13. Pelvis radiograph shows bilateral grade 4 sacroiliitis, entheseal erosion (right ischium), and entheseal ossification (both ischii).

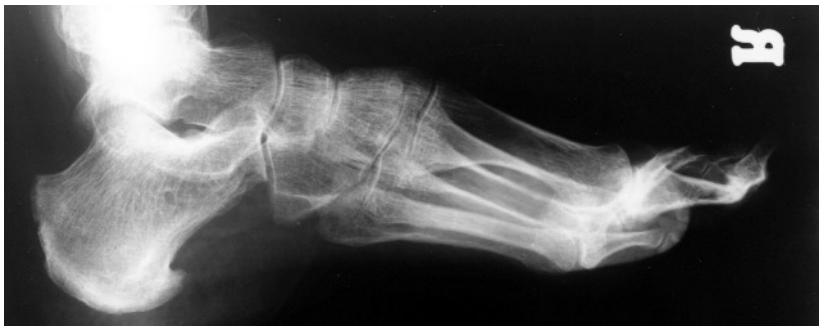

Figure 14. Lateral view of heel shows irregular ossification at the plantar insertion and erosion in the region of the retrocalcaneal bursa close to the Achilles insertion.

in this study. Since agreement beyond chance is likely to be smaller than we have reported in more naturalistic settings, only those features with at least "moderate" reliability should be selected for further study.

On the other hand, the nonstandard presentation of radiographs of varying quality and age in this study does reproduce many of the problems with standardization in the field. Attention to standardization of views and techniques would be expected to increase observer reliability.

A further potential limitation is the heterogeneity of the appearances of each feature in terms of severity. We chose to ignore the severity of each lesion and simply judged whether the lesion was present or absent. However, as the principal purpose of this study was to confirm which features could reasonably be tested for their utility in classification criteria, we felt that severity grading of each feature would be unnecessarily complex. This is similar to the case for the classification criteria for RA, where erosions of any severity are acceptable as one criterion. On the other hand, if radiological features were to be used as outcome measures in PsA, the issue of severity grading becomes much more important.

There are other less common or less specific plain radiographic manifestations of PsA that were not evaluated. These include the "ivory digit," symphisitis, joint erosions other than in DIP joints, zygapophyseal joint involvement, squaring of vertebrae, atlantoaxial subluxation, and manubriosternal or sternoclavicular involvement. We sought

to evaluate features that were sufficiently common and which we thought would be potentially reliable so as to be useful in classification criteria. Also, imaging techniques such as MRI or ultrasound were not addressed in this study.

The diagnostic utility of these features now requires testing in unselected patients with PsA or other inflammatory arthropathies.

ACKNOWLEDGMENT

We thank all the physicians who contributed radiographs, especially Dr. Helena Marzo-Ortega.

REFERENCES

1. Gladman DD. Psoriatic arthritis. Baillieres Clin Rheumatol 1995;9:319-29.
2. Taylor WJ, Helliwell PS. Case definition of psoriatic arthritis [letter]. Lancet 2000;356:2095.
3. Bennett RM. Psoriatic arthritis. In: McCarty DJ, editor. Arthritis and related conditions. 9th ed. Philadelphia: Lea and Febiger; 1979:645.
4. Fournie B, Crognier L, Arnaud C, et al. Proposed classification criteria of psoriatic arthritis. A preliminary study in 260 patients. Rev Rhum Engl Ed 1999;66:446-56.
5. McGonagle D, Conaghan PG, Emery P. Psoriatic arthritis. A unified concept twenty years on. Arthritis Rheum 1999;42:1080-6.
6. Vasey F, Espinoza LR. Psoriatic arthropathy. In: Calin A, editor. Spondyloarthropathies. Orlando, Florida: Grune & Stratton; 1984:151-85.
7. Dale K, Vinje O. Radiology of the spine and sacroiliac joints in ankylosing spondylitis and psoriasis. Acta Radiol Diagn Stockh 1985;26:145-59.
8. Helliwell PS, Hickling P, Wright V. Do the radiological changes of classic ankylosing spondylitis differ from the changes found in the spondylitis associated with inflammatory bowel disease, psoriasis, and reactive arthritis? Ann Rheum Dis 1998;57:135-40.
9. Schumacher TM, Genant HK, Kellet MJ, Mall JC, Fye KH. HLA-B27 associated arthropathies. Radiology 1978;126:289-97.
10. McEwen C, DiTata D, Lingg C, Porini A, Good A, Rankin T. Ankylosing spondylitis and spondylitis accompanying ulcerative colitis, regional enteritis, psoriasis and Reiter's disease. A comparative study. Arthritis Rheum 1971;14:291-318.
11. Avila R, Pugh DG, Slocumb CH, Winkleman RK. Psoriatic arthritis: a roentgenologic study. Radiology 1960;75:691-701.
12. McGonagle D, Khan MA, Marzo-Ortega H, O'Connor P, Gibbon W, Emery P. Enthesitis in spondyloarthropathy. Curr Opin Rheumatol 1999;11:244-50.
13. Baker H, Golding DN, Thompson M. Psoriasis and arthritis. Ann Intern Med 1963;58:909-25.
14. Wright V. Psoriasis and arthritis. A study of the radiographic appearances. Br J Radiol 1957;30:113-9.
15. Blau RH, Kaufman RL. Erosive and subluxing cervical spine disease in patients with psoriatic arthritis. J Rheumatol 1987;14:111-7.
16. Marsal S, Armadans-Gil L, Martinez M, Gallardo D, Ribera A, Lience E. Clinical, radiographic and HLA associations as markers for different patterns of psoriatic arthritis. Rheumatology Oxford 1999;38:332-7.
17. Fiumicelli A, Bruni L. Psoriatic arthritis. Rays 1989;14:237-45.
18. Torre Alonso JC, Rodriguez Perez A, Arribas Castrillo JM, Ballina Garcia J, Riestra Noriega JL, Lopez Larrea C. Psoriatic arthritis: a clinical, immunological and radiological study of 180 patients. Br J Rheumatol 1991;30:245-50.
19. McRae IF, Haslock DI, Wright V. Grading of films for sacroiliitis in population studies. Ann Rheum Dis 1971;30:58-66.
20. Spoorenberg A, de Vlam K, van der Heijde D, et al. Radiological scoring methods in ankylosing spondylitis: reliability and sensitivity to change over one year. J Rheumatol 1999;26:997-1002.
21. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993;46:423-9.
22. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. J Clin Epidemiol 1990;43:551-8.
23. Mackinnon A. A spreadsheet for the calculation of comprehensive statistics for the assessment of diagnostic tests and inter-rater agreement. Comp Biol Med 2000;30:127-34.
24. Resnick D, Niwayama G. Psoriatic arthritis. In: Resnick D, Niwayama G, editors. Diagnosis of bone and joint disorders. 2nd ed. Philadelphia: W.B. Saunders Co.; 1988:1171-98.
25. Cheeseborough MJ. Osteolysis and psoriasis. Clin Exp Dermatol 1979;4:341-4.
26. Sharp JT, Lidsky MD, Collins LC, Moreland J. Methods of scoring the progression of radiologic changes in rheumatoid arthritis. Correlation of radiologic, clinical and laboratory abnormalities. Arthritis Rheum 1971;14:706-20.
27. Kerr R, Resnick D. Radiology of the seronegative spondyloarthropathies. Clin Rheum Dis 1985;11:113-46.
28. Porter GG. Psoriatic arthritis. Plain radiology and other imaging techniques. Baillieres Clin Rheumatol 1994;8:465-82.
29. Forrester DM, Kirkpatrick J. Periostitis and pseudoperiostitis. Radiology 1976;118:597-601.
30. Resnick D, Broderick TW. Bony proliferation of the terminal toe phalanx in psoriasis: the "ivory" phalanx. J Can Assoc Radiol 1977;28:187-9.
31. Bywaters EGL, Dixon ASJ. Paravertebral ossification in psoriatic arthritis. Ann Rheum Dis 1965;24:313-31.
32. Sundaram M, Patton JT. Paravertebral ossification in psoriasis and Reiter's disease. Br J Radiol 1975;48:628-33.
33. Killebrew K, Gold RH, Sholkoff SD. Psoriatic spondylitis. Radiology 1973;108:9-16.
34. Cawley MID, Chalmers TM, Kellgren JH, Ball J. Destructive lesions of vertebral bodies in ankylosing spondylitis. Ann Rheum Dis 1972;31:345-58.
35. Andersson O. Rontgenbilden vid spondylarthritis ankylopoetica. Nord Med Tidskr 1937;14:2000.
36. Salvarani C, Macchioni P, Cremonesi T, et al. The cervical spine in patients with psoriatic arthritis: a clinical, radiological and immunogenetic study. Ann Rheum Dis 1992;51:73-7.
37. Jenkinson T, Armas J, Evison G, Cohen M, Lovell C, McHugh NJ. The cervical spine in psoriatic arthritis. Br J Rheumatol 1994;33:255-9.
38. Lambert JR, Wright V. Psoriatic spondylitis: a clinical and radiological description of the spine in psoriatic arthritis. QJM 1977;46:411-25.
39. Gofton JP. Report from the Subcommittee on Diagnostic Criteria for Ankylosing Spondylitis. In: Bennett PH, Wood HN, editors. Population studies of the rheumatic diseases; 1966. New York: Excerpta Medica; 1966:314-6, 456-7.
40. van der Linden S, Valkenberg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modifications of the New York criteria. Arthritis Rheum 1984;27:361-8.
41. McGonagle D, Gibbon W, Emery P. Classification of inflammatory arthritis by enthesitis. Lancet 1998;352:1137-40.
42. Ruhoy MK, Schweitzer ME, Resnick D. Spondyloarthropathies: enthesopathy. In: Klippel JH, Dieppe PA, editors. Rheumatology. London: Mosby; 1998.
43. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
44. Scapa R. Discovertebral erosions and destruction in psoriatic

arthritis. J Rheumatol 2000;27:975-8.
45. Brower AC. Spondyloarthropathies: Imaging. In: Klippel JH, Dieppe PA, editors. Rheumatology. 2nd ed. London: Mosby; 1998.
46. Resnick D, Niwayama G. Ankylosing spondylitis. In: Resnick D, Niwayama G, editors. Diagnosis of bone and joint disorders. 2nd ed. Philadelphia: W.B. Saunders; 1988:1118-9.
47. Battistone MJ, Manaster BJ, Reda DJ, Clegg DO. The prevalence of sacroiliitis in psoriatic arthritis: new perspectives from a large, multicenter cohort. Skeletal Radiol 1999;28:196-201.
48. Harvie JN, Lester RS, Little AH. Sacroiliitis in severe psoriasis. Am J Roentgenol 1976;127:579-84.
49. Peterson CC, Silbiger ML. Reiter's syndrome and psoriatic arthritis — their roentgen spectra and some interesting similarities. AJR Am J Roentgenol 1967;101:860-71.
50. Resnick D, Niwayama G. Entheses and enthesopathy — anatomical, pathological and radiological correlation. Radiology 1983;146:1-9.
51. Resnick D, Niwayama G. Bony proliferation in "rheumatoid variant" disorders. AJR Am J Roentgenol 1977;129:275-8.
52. Martel W, Stuck KJ, Dworin AM, Hylland RG. Erosive osteoarthritis and psoriatic arthritis: a radiologic comparison in the hand, wrist and foot. Am J Roentgenol 1980;134:125-135.
53. Meany TF, Hays RA. Roentgen manifestations of psoriatic arthritis. Radiology 1957;681:403-7.
54. Sherman MS. Psoriatic arthritis: observations on the clinical, roentgenographic and pathological changes. J Bone Joint Surg Am 1952;34:831-52.
55. Eisenstadt HB, Eggers GWN. Arthritis mutilans (doigt, main, pied en lorgnette). J Bone Joint Surg Am 1955;37:337-46.