Validation of Definitions for Active Inflammatory Lesions Detected by Magnetic Resonance Imaging in the Spine of Patients with Spondyloarthritis

SUSANNE J. PEDERSEN, MIKKEL ØSTERGAARD, PRAVEENA CHIOWCHANWISAWAKIT, ROBERT G.W. LAMBERT, and WALTER P. MAKSYMOWYCH

ABSTRACT. ***Objective.*** Assessment of active inflammatory features of spondyloarthritis (SpA) on magnetic resonance imaging (MRI) is of diagnostic and prognostic significance. Further study requires standardization and application of rigorous definitions of the spinal changes observed on MRI. The Canada/Denmark MRI working group has developed a comprehensive list of definitions for active inflammatory spinal lesions. We aimed to conduct a systematic evaluation of the frequency and reliability of detection of active inflammatory lesions as defined by the working group.

Methods. Four readers independently recorded dichotomously (present/absent) active inflammatory lesions from lower C2 to the upper sacrum of the spine on STIR MRI scans of 20 patients with ankylosing spondylitis. All lesions were recorded at individual spinal levels except for facet joint lesions, which were recorded according to affected spinal segments. Prior to the exercise, a series of reference images were developed in which active inflammatory lesions were assigned by consensus. Frequency data were analyzed descriptively, while reliability was assessed by Cohen's unweighted kappa and percentage agreement.

Results. Interobserver reliability of vertebral corner inflammatory lesions (CIL) varied substantially between reader pairs and between spinal segments. It was overall less than adequate (kappa < 0.60) for most reader pairs, although the most experienced reader pair achieved good reliability (kappa = 0.68). Reliability was fair to good for lateral segment inflammatory lesions (LIL) and non-corner inflammatory lesions (NIL) (mean kappa for experienced reader pair 0.58 and 0.66, respectively).

Conclusion. Detection of CIL, LIL, and NIL was only satisfactory with the most experienced MRI reader pair. Despite rigorous standardization of definitions, detection of active inflammatory lesions is difficult and requires substantial calibration. (J Rheumatol 2009;36 Suppl 84:35-38; doi:10.3899/jrheum.090618)

Key Indexing Terms:
SPONDYLOARTHRITIS
INFLAMMATION
MAGNETIC RESONANCE IMAGING
ACTIVE INFLAMMATORY LESIONS

From the Department of Rheumatology, Copenhagen University Hospital at Gentofte; Department of Radiology, Copenhagen University Hospital at Herlev; Departments of Rheumatology, Copenhagen University Hospitals at Hvidovre, Glostrup, and Gentofte, Copenhagen, Denmark; Department of Medicine, University of Alberta, and Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand; and Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, Alberta, Canada.

We are grateful for the support provided through the Nathional Research Initiative from The Arthritis Society.

W.P. Maksymowych is a Scientist of the Alberta Heritage Foundation for Medical Research

S.J. Pedersen, MD, Rheumatology Fellow, Department of Rheumatology, Copenhagen University Hospital at Gentofte, Department of Radiology, Copenhagen University Hospital at Herlev; M. Østergaard, MD, PhD, DMSc, Professor, Departments of Rheumatology, Copenhagen University Hospitals at Hvidovre and Glostrup; P. Chiowchanwisawakit, MD, Rheumatologist, Department of Medicine, University of Alberta, Faculty of Medicine, Siriraj Hospital, Mahidol University; R.G.W. Lambert, FRCPC, Professor, Department of Radiology and Diagnostic Imaging, University of Alberta; W.P. Maksymowych, FRCPC, Professor, Department of Medicine, University of Alberta.

Address correspondence to Dr. W. Maksymowych, 562 Heritage Medical Research Building, University of Alberta, Edmonton, Alberta T6G 2S2, Canada. E-mail: walter.maksymowych@ualberta.ca

Magnetic resonance imaging (MRI) is the most sensitive imaging modality for the detection of inflammatory lesions in the spine of patients with spondyloarthritis (SpA)[1]. Systematic evaluation has also shown that lesions occur frequently in anterior, posterior, and lateral aspects of the vertebral bodies, and recent recommendations emphasize the importance of diagnostic MRI study protocols that enhance detailed scrutiny of all regions of the spine that may be neglected (particularly in patients with scoliosis) using conventional imaging protocols[2,3]. The diagnostic utility and prognostic significance of these lesions remain an area of active study, and further progress depends on the development of standardized definitions that describe these lesions in a manner that allows reliable detection not only by clinical researchers but also practising clinicians and radiologists in all spheres of practice.

The Canada-Denmark MRI working group has developed a list of definitions that describe active inflammatory lesions in the spine of patients with SpA. These are based on the increased signal observed on short-tau inversion recovery (STIR) or T2-weighted fat suppressed (T2FS) sagittal sequences of the spine. Active inflammatory lesions in the anterior spine on central sagittal slices

that include the spinal canal have been termed anterior and posterior vertebral corner inflammatory lesions (aCIL and pCIL), which are the imaging counterpart of the histopathologically defined Romanus lesion, and non-corner inflammatory lesions (NIL) adjacent to vertebral endplates that are often termed spondylodiscitis. A lesion may occasionally extend across the whole length of the vertebral endplate, in which case it is termed a massive inflammatory lesion (MIL). Active inflammatory lesions adjacent to the vertebral endplate observed in lateral sagittal slices that do not include the spinal canal have been termed lateral inflammatory lesions (LIL). Finally, we have also defined active inflammatory lesions in the posterior spine adjacent to a facet joint, the facet inflammatory lesion (FIL), or in another posterior element (PIL). These are all clearly depicted in reference images developed by the working group that include lesions considered to be at the limit of detection[4].

We tested the list of standardized definitions for active inflammatory lesions in unselected patients with ankylosing spondylitis (AS) to determine if they can be used reliably.

METHODS

Patients. We studied 20 patients recruited to an observational cohort of patients with SpA who were systematically assessed every 6 months according to clinical, laboratory, and imaging outcomes. Patients fulfilled the modified New York criteria for ankylosing spondylitis (AS)[5]. There were 14 men of mean (SD) age 43.8 (14.6) years and mean (SD) disease duration of 17.1 (11.3) years. Mean Bath AS Disease Activity Index (BASDAI) was 4.3 (range 3.1–5.9) and mean Bath AS Functional Index (BASFI) was 4.0 (range 2.2–6.4). The study was approved by the University of Alberta ethics committee, and all patients provided informed consent.

Magnetic resonance imaging. All MRI of the spine were performed with 1.5 Tesla systems (Siemens, Erlangen, Germany) using appropriate surface coils. Sagittal sequences were obtained with 3–4 mm slice thickness and 11–15 slices were acquired. Sequence parameters were: 1. T1-weighted spin echo (TR 517–618 ms, TE 13 ms); 2. STIR (TR 3000–3170 ms, TI 140 ms, TE 38–61 ms). Field of view was 380–400 mm and matrix was 512 x 256 pixels. The spine was imaged in 2 parts: 1. upper half comprising the entire cervical and most of the thoracic spine; 2. lower half comprising the lower portion of the thoracic spine and entire lumbar spine. The specific MRI parameters for acquiring spine images are provided on our website (www.arthritisdoctor.ca).

MRI reading exercises. A unique MRI study number was allocated to each patient, ensuring blinding to all patient demographics. Allocation was done by a technologist unconnected with the study using computer generated random numbers. Assessment was performed on a 3-monitor review station by 4 readers using computer software that is optimal for this type of review (Merge Healthcare eFilm, Milwaukee, WI, USA) and viewing conditions were standardized. Individual subjects were identified only by the MRI study number, and scans were read by 2 experienced MRI readers (WPM, MØ) and 2 rheumatology fellows (SJP, PC) who had participated in the group's deliberations and MRI reading exercises since its inception. MR images for each patient were evaluated at 25 spinal levels from C1 to S1. We recorded inflammatory signal (dichotomous yes/no) on STIR images at each spinal level for the following definitions of active inflammatory lesions proposed by the Canada-Denmark MRI working group: aCIL, pCIL, MIL, NIL, LIL. Type B CIL, where the inflammatory lesion itself does not extend to the corner but extends to both the adjacent vertebral endplate and the anterior or posterior cortex of the vertebral body[4], were also recorded. FIL and PIL were recorded only as being present/absent at a spinal segmental level because deliberations during selection of reference images led the group to conclude that reliability assessed at the level of individual spinal levels would be poor. Calibration of readers was accomplished during 3 video teleconferences aimed at the selection of reference images by consensus that included examples of each of these lesions, including images demonstrating lesions at the threshold of detection. These reference images were available for reference purposes during the reading exercise.

Statistics. The distribution of active inflammatory lesions according to patient, spinal segment, and individual vertebral level was analyzed descriptively. Data provided here represent the mean scores for the 4 readers.

Analysis of interobserver reliability for detection of inflammatory changes addressed reliability according to detection of lesions (yes/no) at individual spinal levels as well as at spinal segments.

Interobserver reliability was assessed using kappa statistics and percentage agreement. Kappa values of > 0.75, 0.40–0.75, and < 0.40 were designated as representing excellent, fair to good, and poor reliability, respectively.

RESULTS

Descriptive data. The majority of patients had an inflammatory lesion recorded in at least one spinal level (92.3%). The lesion seen most frequently in patients was an aCIL (66.7%), followed by LIL (57.5%), pCIL (51.5%), PIL (47.5%), FIL (42.3%), and NIL (28.5%). The mean (SD) number of lesions per patient was as follows: aCIL 3.6 (2.1), LIL 2.7 (1.2), pCIL 1.6 (0.9), NIL

0.5 (0.3). The thoracic and lumbar spine were most frequently affected by inflammation, this being evident at almost 20% of spinal levels per patient (Table 1). CIL were more frequent at lumbar spinal levels, while LIL were more frequent at thoracic spinal levels.

Reliability data. Interobserver reliability of CIL varied substantially between reader pairs and between spinal segments (Table 2). It was overall less than adequate for most reader pairs, although the most experienced reader pair (readers 1 and 4) achieved good reliability for both aCIL and pCIL in most segments of the spine. Similarly, interobserver reliability varied for the detection of NIL and LIL and was highest for the 2 experienced readers (kappa = 0.66 and 0.58, respectively; Table 3). Although kappa values were overall lower for NIL than LIL, percentage agreement was higher for NIL, reflecting the low frequency of these lesions.

DISCUSSION

Our study demonstrates that good reliability for active inflammatory lesions is difficult to attain, the highest scores being recorded in the thoracic spine and for lateral inflammatory lesions. Better reliability was attained by the most experienced MRI readers, indicating that in addition to standardization of definitions, further training and calibration between readers is necessary.

After review of these results the conclusion of the group was that further modification of the definitions was unlikely to improve reliability. There are several factors that limit reliability related to the approach to imaging of the spine and inherent limitations of MRI. Many lesions are small and are depicted by only a few pixels of increased signal intensity on STIR sequences. Moreover, the large field of view that results from scanning the entire spine in 2 halves means that the resolution of cervical vertebrae is poor, and inflammatory lesions may

Table 1. Mean percentage (range) of spinal levels per patient with inflammatory lesions in the cervical, thoracic, lumbar, and total spine as assessed by 4 readers in 20 patients with SpA.

	Any Inflammatory Lesion	aCIL	pCIL	MIL	NIL	LIL
C spine	7.6 (2.3–11.4)	3.3 (0.5–6.4)	3.3 (2.3–4.5)	1.1 (0–3.2)	0.2 (0–0.9)	0.1 (0–0.5)
T spine	19.7 (13.5–24.4)	8.6 (5.4–13.1)	5.1 (2.1–8.5)	1.3 (0.8–1.7)	1.6 (1–2.5)	8.9 (5.6–10.6)
L spine	19.3 (15–26.8)	13.4 (8.2–20.5)	4.8 (0.9–7.7)	0.3 (0–0.9)	1.4 (0–4.1)	4.8 (3.1–5.9)
Total	16.7 (11.2–21.8)	8.5 (4.9–13.3)	4.6 (2.1–7.4)	1.0 (0.4–1.6)	1.2 (0.7–2.5)	5.8 (4.2–6.4)

aCIL: anterior vertebral corner inflammatory lesion; pCIL: posterior vertebral corner inflammatory lesion; NIL: non-corner inflammatory lesion; MIL: massive inflammatory lesion; LIL: lateral inflammatory lesion; FIL: facet inflammatory lesion; PIL: posterior inflammatory lesion.

Table 2. Interobserver reliability (pecentage agreement/kappa values) of active inflammatory vertebral corner (anterior and posterior) lesions in the spine as assessed by 4 readers in 20 patients with SpA.

Interreader Pairs	C Spine		T Spine		L Spine		Total	
	aCIL	pCIL	aCIL	pCIL	aCIL	pCIL	aCIL	pCIL
1 and 2	97.3/0.24	96.4/0.32	93.8/0.50	95.0/0.18	87.7/0.39	95.9/0.30	93.2/0.45	95.5/0.23
1 and 3	93.2/0.25	97.7/0.69	91.9/0.57	92.3/0.25	91.9/0.57	93.2/0.20	90.7/0.52	93.8/0.32
1 and 4	97.3/0.56	99.1/0.85	96.3/0.74	95.2/0.36	91.4/0.61	95.5/0.27	95.3/0.68	96.2/0.45
2 and 3	94.1/0.13	95.9/0.38	89.8/0.40	94.0/0.50	85.9/0.44	92.7/0.39	89.9/0.40	94.1/0.45
2 and 4	97.3/0.24	96.4/0.32	94.2/0.53	97.3/0.71	91.8/0.53	99.5/0.95	94.3/0.52	97.6/0.72
3 and 4	93.2/0.25	97.7/0.70	91.0/0.53	94.6/0.59	85.0/0.44	93.2/0.45	90.1/0.47	95.0/0.57
Mean	95.4/0.28	97.2/0.54	92.8/0.55	94.7/0.43	88.9/0.50	95/0.43	92.3/0.51	95.4/0.46

aCIL: anterior vertebral corner inflammatory lesion; pCIL: posterior vertebral corner inflammatory lesion; NIL: non-corner inflammatory lesion; MIL: massive inflammatory lesion; LIL: lateral inflammatory lesion; FIL: facet inflammatory lesion; PIL: posterior inflammatory lesion.

Table 3. Interobserver reliability (percentage agreement/kappa values) of acute non-corner vertebral (NIL) and lateral inflammatory (LIL) lesions in the spine as assessed by 4 readers in 20 patients with SpA.

Interreader Pairs	NIL	LIL
1 and 2	99.3/0.50	93.7/0.38
1 and 3	98.2/0.41	95.0/0.58
1 and 4	99.5/0.66	95.1/0.58
2 and 3	97.5/0.20	93.3/0.33
2 and 4	99.0/0.40	95.8/0.57
3 and 4	98.0/0.43	95.1/0.58
Mean	98.6/.043	94.7/0.50

be very small and not reliably detected unless signal intensity is high. In the lumbar spine, signal artefact over the anterior portion of the vertebrae seen on the STIR sequence due to blood flow in the great vessels may cause confusion with the increased signal due to inflammation, especially if signal intensity is low. This is less of a concern in the lateral aspect of the vertebral bodies. Although the mean kappa value for detection of LIL and CIL was comparable, the distribution of kappa values was lower for LIL than for CIL. The apparently low kappa value for NIL does not necessarily indicate that reliability is poor in view of the very low frequency of these lesions, and percentage agreement was higher than for LIL. Detection was sometimes difficult when lesions were small and/or difficult to distinguish from discal abnormalities in the setting of irregularity and/or erosion of the vertebral endplate.

Based on these observations, it is our recommendation that assessment of inflammatory lesions for studies of diagnostic and prognostic utility should be conducted by 2 calibrated readers evaluating images independently, in contrast to studies where only a single reader has assessed the images or where several readers have generated imaging data by consensus. Concordant imaging data from both readers should then constitute the principal variable for analysis, as demonstrated in a recent prospective study assessing the prognostic value of CIL in the development of new syndesmophytes[6].

REFERENCES

1. Maksymowych WP, Lambert RG. Magnetic resonance imaging for spondyloarthritis — avoiding the minefield. J Rheumatol 2007;34:259-65.
2. Rennie WJ, Dhillon SS, Conner-Spady B, Maksymowych WP, Lambert RG. Magnetic resonance imaging assessment of spinal inflammation in ankylosing spondylitis: standard clinical protocols may omit inflammatory lesions in thoracic vertebrae. Arthritis Rheum 2009;61:1187-93.
3. Crowther SM, Lambert RGW, Dhillon SS, Maksymowych WP. High frequency of inflammatory lesions in the posterior structures of the spine in patients with ankylosing spondylitis (AS): A systematic evaluation by MRI [abstract]. Arthritis Rheum 2006;54 Suppl:S793.
4. Lambert RGW, Pedersen SJ, Maksymowych WP, Chiowchanwisawakit P, Østergaard M. Active inflammatory lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis – definitions, assessment system, and reference image set. J Rheumatol 2009;36 Suppl 84:3-17.
5. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361-8.
6. Maksymowych WP, Chiowchanwisawakit P, Clare T, Pedersen S, Ostergaard M, Lambert RGW. Inflammatory lesions of the spine on MRI predict the development of new syndesmophytes in ankylosing spondylitis: Evidence for coupling between inflammation and new bone formation. Arthritis Rheum 2009;60:93-102.