

Low-dose Infliximab (3 mg/kg) Significantly Reduces Spinal Inflammation on Magnetic Resonance Imaging in Patients with Ankylosing Spondylitis: A Randomized Placebo-controlled Study

WALTER P. MAKSYMOWYCH, DAVID SALONEN, ROBERT D. INMAN, PROTON RAHMAN, and ROBERT G.W. LAMBERT, for the CANDLE Study Group

ABSTRACT. *Objective.* To evaluate the influence of low-dose infliximab (IFX) on spinal inflammation scored by magnetic resonance imaging (MRI). The dose recommended for rheumatoid arthritis (3 mg/kg) is also clinically effective for ankylosing spondylitis (AS), although effects on spinal inflammation as defined by MRI have yet to be described in a placebo-controlled trial.

Methods. In a 12-week double-blind period, patients were randomized 1:1 to receive either IFX 3 mg/kg at 0, 2, and 6 weeks, or placebo. Spinal inflammation in discovevertebral units (DVU) was measured by the Spondyloarthritis Research Consortium of Canada (SPARCC) MRI Index at baseline and 12 weeks by 3 readers blinded to timepoint and treatment allocation. We also compared reliability and discrimination of the SPARCC MRI index based on evaluation of the entire spine (23 DVU score) compared to assessment of only the 6 most severely affected DVU (6 DVU score).

Results. At Week 12, patients treated with IFX experienced mean reductions of 55.1% and 57.2% in the 6 DVU and 23 DVU SPARCC scores, respectively, compared with a mean increase of 5.8% and decrease of 3.4% in 6 DVU and 23 DVU scores, respectively, for patients taking placebo ($p < 0.001$). A large treatment effect (Guyatt's effect size ≥ 1.7) and high reliability was evident and comparable between 6 DVU and 23 DVU scoring methods.

Conclusion. Treatment with low-dose IFX leads to a large treatment effect on spinal inflammation as measured by MRI. Scoring for inflammation of only the most severely affected regions of the spine by MRI is comparable to assessment of the entire spine. (First Release May 1 2010; J Rheumatol 2010;37:1728–34; doi:10.3899/jrheum.091043)

Key Indexing Terms:

INFLIXIMAB
MAGNETIC RESONANCE IMAGING

SPINAL INFLAMMATION
ANKYLOSING SPONDYLITIS

Ankylosing spondylitis (AS) is a chronic inflammatory disorder of the axial skeleton that may result in substantial disability and reduction in quality of life. Two major advances in the evaluation and management of this disease have been the introduction of magnetic resonance imaging (MRI) for the objective assessment of inflammation, and the substan-

tial efficacy of anti-tumor necrosis factor- α (anti-TNF) agents. Inflammation on MRI is detected using fat-suppressed sequences, such as short-tau inversion recovery (STIR), which suppresses the signal from bone marrow fat, allowing visualization of the free water signal associated with underlying bone marrow edema. Inflammation in the spine is then readily detected within bone marrow at anterior or vertebral corners and adjacent to the vertebral endplate remote from the vertebral corner. The thoracic spine is most frequently affected. Imaging of the spine is conducted in the sagittal plane of view and the spine is imaged in cervicothoracic and thoracolumbar portions. A particular advantage of MRI for assessment of spinal inflammation is its ability to visualize lesions in 3 planes of view by assessing consecutive sagittal slices through the spine. Systematic study has also shown that it is important to assess lesions in the lateral segments of the spine where inflammation is particularly frequent within and adjacent to the costovertebral joints¹.

Several scoring systems have been developed to permit quantification of the extent of inflammation in the spine,

From the Department of Medicine and the Department of Radiology, University of Alberta, Edmonton, Alberta; Department of Medicine, University of Toronto, Toronto, Ontario; and Department of Medicine, Memorial University, St. John's, Newfoundland, Canada.

Supported by Schering-Plough Canada. Dr. Maksymowych is a Scientist of the Alberta Heritage Foundation for Medical Research.

W.P. Maksymowych, MB, FRCPC, Department of Medicine, University of Alberta; D. Salonen, MD, FRCPC, Radiologist, University of Toronto; R.D. Inman, MD, FRCPC, Professor, Department of Medicine, University of Toronto; P. Rahman, MD, FRCPC, Professor, Department of Medicine, Memorial University; R.G.W. Lambert, MB, FRCPC, Professor, Department of Radiology, University of Alberta.

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

Accepted for publication February 12, 2010.

particularly for the objective evaluation of therapeutic agents in clinical trials². These differ in several respects, a principal difference being whether the entire spine or only a limited number of the most severely affected spinal units are assessed^{3,4}. Each method scores inflammation within a single spinal unit, termed a discovertebral unit (DVU), an area defined by 2 imaginary horizontal lines through the middle of adjacent vertebrae on a sagittal plane of view that contains the disc and the 2 vertebral endplates with adjacent bone marrow. Scoring the entire spine is laborious and may result in scoring of artifacts and/or lesions with limited anatomical resolution, which may reduce reliability and discrimination⁵. The Spondyloarthritis Research Consortium of Canada (SPARCC) MRI index was developed to specifically address these concerns by quantifying lesions in only the 6 most severely affected DVU⁶. This number was chosen because prior study had demonstrated that the median number of affected DVU per patient with AS was 3–4⁷. Subsequent research showed that reliability and responsiveness was comparable when either 6 DVU or all 23 spinal DVU were assessed, although discrimination was not examined⁴.

Several anti-TNF- α agents have now been shown to be efficacious in patients with AS. These include infliximab (IFX) when administered in a dose of 5 mg/kg every 6 to 8 weeks^{8,9}. Open-label studies have shown that a lower dose of 3 mg/kg every 8 weeks may also be efficacious^{10,11}. A Canadian randomized placebo-controlled trial study has now verified the safety and efficacy of low-dose IFX (3 mg/kg) for AS¹². Discordance of clinical and radiographic outcomes has been noted in previous studies and it was important to evaluate the effect of low-dose IFX on MRI outcomes as well as clinically defined outcomes. MRI was used at 2 institutions participating in our study to evaluate the effect of low-dose IFX on objective features of spinal inflammation. This also allowed comparative study of the relative performance of a scoring method limited to only the most severely affected regions of the spine compared to mandatory assessment of the entire spine.

MATERIALS AND METHODS

Patients. Patients were adults (≥ 18 years of age) with a diagnosis of AS defined by the modified New York criteria¹³ who were not responsive to or were intolerant of ≥ 1 nonsteroidal antiinflammatory drug (NSAID). Patients who had failed 1 or more disease-modifying antirheumatic drugs (DMARD; e.g., methotrexate, sulfasalazine) were also allowed to enroll. Active AS at baseline was defined by a Bath AS Disease Activity Index (BASDAI)¹⁴ score ≥ 4 . Patients could continue sulfasalazine (≤ 3 g/day), methotrexate (≤ 25 mg/wk), hydroxychloroquine (≤ 400 mg/day), prednisone, and/or prednisone equivalents (≤ 10 mg/day), and/or NSAID as long as these doses had remained stable before baseline (14 days for NSAID, 30 days for DMARD/steroid). All patients were evaluated for latent tuberculosis infection at baseline using a purified protein derivative skin test and chest radiograph; patients with evidence of latent tuberculosis infection were allowed to participate if a documented history of antituberculous treatment was available or if prophylactic antituberculous treatment was initiated before the first dose of IFX. Patients were excluded from the

study if they had a history of chronic/recurrent infectious disease, including tuberculosis, hepatitis B, or human immunodeficiency virus, and/or a diagnosis of malignancy or lymphoproliferative disease currently or within the past 5 years. Patients who had previously received TNF- α antagonist therapy or 1 or more intraarticular joint injections with corticosteroids within 4 weeks before the baseline visit were excluded. Patients with radiologic evidence of total spinal ankylosis (bamboo spine) were excluded.

Our study was approved by an independent ethics committee at the 2 study centers that conducted MRI examination, the University of Alberta and the University of Toronto. Our study was performed in accord with the ethical principles that originate in the Declaration of Helsinki. Written, informed consent was obtained from each patient.

Study design. This Phase IIb, randomized, multicenter, double-blind, placebo-controlled study was conducted at 8 investigational centers in Canada, and all patients enrolled at 2 sites (University of Alberta, University of Toronto) were invited to have MRI examination at baseline and Week 12. Patients were randomly assigned in a 1:1 ratio to receive infusions of either placebo or IFX 3 mg/kg at Weeks 0, 2, and 6, with evaluation for primary and secondary endpoints at Week 12. The primary endpoint of this study was the proportion of patients achieving a 20% improvement in AS assessment in AS (ASAS) International Working Group criteria (ASAS20) after 12 weeks of treatment¹⁵. Secondary endpoints at Week 12 included the change from baseline in BASDAI, Bath AS Functional Index¹⁶ (BASFI), C-reactive protein (CRP; mg/l), and erythrocyte sedimentation rate (ESR), as well as ASAS40, ASAS50, ASAS70, and ASAS 5/6¹⁷.

An MRI of the spine was performed at baseline and Week 12 using appropriate surface coils for systems operating at 1.5 Tesla. STIR sequences were obtained in sagittal orientation, with 12 to 15 slices, 3–4 mm thick, acquired using the following measurements: repetition time 2720–3170 ms; echo time 38–61 ms; time to inversion 140 ms. Imaging of the spine was divided into 2 parts: (1) the entire cervical spine and most of the thoracic spine, and (2) the lower portion of the thoracic spine and the entire lumbar spine. T1 spin-echo images of the entire spine were also obtained for use as anatomical references.

Scoring of MRI lesions. The SPARCC scoring method is based on an abnormal increased signal on the STIR sequence, representing bone marrow edema (defined as an increased concentration of “free water” relating to a bone lesion). Examples of the scoring method for the spine have been published^{4,6}. Details of each scoring method are provided at the website www.arthritisdoctor.ca. The scoring method designed for use in clinical trials requires that the entire spine be assessed first and the 6 most severely affected DVU then selected for formal scoring of bone marrow edema. A single DVU has a scoring range of 0–18, so this results in a total scoring range of 0–108. However, it is also possible to evaluate and score bone marrow edema at all 23 DVU.

Viewing conditions. All scans were reviewed on workstations with 2–4 large screens and image-manipulation software. This system permitted simultaneous display of all MRI sequences (T1 and STIR for upper and lower spine for both timepoints) at original (life-size) dimensions. Scores were recorded electronically on an additional screen by entering data into an online scoring sheet that comprised a schematic of 3 DVU divided into quadrants to allow scoring of each DVU in 3 consecutive sagittal slices; the reader was able to see all scores and all scans simultaneously before committing to the final score. Each image was rated by 3 independent readers to allow calculation of interreader scoring variability. Two of the readers were qualified, fellowship-trained musculoskeletal radiologists and one is a rheumatologist who is a codeveloper of the SPARCC method. Readers were blinded to patients’ identities, treatments, and imaging timepoints.

Statistical analyses. Because the SPARCC scores were secondary endpoints of the overall study, there was no *a priori* power calculation for the MRI subanalysis. All statistical tests were 2-sided and comparisons were performed with $\alpha = 0.05$. The changes in SPARCC scores from baseline to Week 12 were compared between the 2 treatment groups using an analysis

of covariance (ANCOVA) model with the baseline score as a covariate. Interreader reliability of baseline and 12-week scores as well as the change from baseline to 12 weeks were determined by calculating intraclass correlation coefficients (ICC) using an ANOVA model with SPARCC scores as the dependent variable, and patient and reader (fixed factor) as independent variables. An ICC value > 0.6 , > 0.8 , and > 0.9 indicates good, very good, and excellent reproducibility, respectively. The association between change in selected clinical variables and the change in SPARCC scores from baseline to Week 12 was calculated using an ANCOVA model adjusted for baseline SPARCC score, treatment, and change in each of the following measures: BASDAI, BASFI, ESR, and CRP. Guyatt's effect size for the change in SPARCC scores from baseline to Week 12 was calculated by dividing the mean change in the IFX group by the SD of the change in the placebo group. Effect sizes of at least 0.2, 0.5, and 0.8 are considered small, moderate, and large, respectively¹⁸.

RESULTS

Patient characteristics. A total of 36 patients were enrolled at the 2 study sites that participated in the MRI portion of the study, of which 18 were randomized to placebo and 18 to IFX. Two patients in each group withdrew prior to 12 weeks, so that 16 in each treatment group had baseline and 12-week MRI scans. Demographic and baseline clinical characteristics were similar in both treatment groups and consistent with a typical AS population that is refractory to standard therapy (Table 1).

Clinical outcomes. The mean (SD) change in BASDAI was -2.3 (1.9) in patients treated with IFX, and -0.8 (2.2) in patients given placebo ($p = 0.05$). Five of 16 (31.3%) patients who received IFX had a BASDAI 50 response, compared to 1 of 16 (6.3%) patients receiving placebo ($p =$ nonsignificant). ASAS20, 40, 50, and 70 responses were noted in 68.8%, 56.3%, 50%, and 12.5% of patients receiving IFX, and 37.5%, 12.5%, 12.5%, and 0% of those receiving placebo, respectively ($p = 0.08$, 0.009, 0.02; $p =$ nonsignificant).

SPARCC MRI spinal inflammation scores. There were no

significant treatment group differences in mean SPARCC MRI 6 DVU or 23 DVU scores at baseline (Tables 2 and 3). No spinal inflammation was evident at baseline in 2 patients, both of whom received IFX, as confirmed by all 3 readers. Five patients (4 IFX, 1 placebo) had no spinal inflammation at baseline as recorded by at least 1 reader. The 6 DVU score was equivalent to the 23 DVU score in 65.6% of readers and would have identified the entire inflammation in the anterior spine in 21 of the 32 patients. The 23 DVU score was greater than the 6 DVU score by a mean (SD) of 5.8 (12.4) and a median (range) of 0 (0–55).

At Week 12, patients treated with IFX experienced a significantly ($p \leq 0.002$) greater reduction in the mean spinal 6 DVU and 23 DVU SPARCC score, compared with patients taking placebo, as recorded by all 3 readers. A reduction in score was recorded by each of the 3 readers in every patient treated with IFX where inflammation was recorded at baseline. A significant reduction in 6 DVU score was also recorded in ASAS20 nonresponders when patients treated with IFX [$n = 7$; mean (SD) change in 6 DVU score = -14.4 (13.2)] were compared to patients taking placebo [$n = 10$; mean (SD) change in 6 DVU score = $+1.7$ (4.8); $p = 0.02$]. Patients treated with IFX experienced a 55.1% mean reduction in the 6 DVU SPARCC score, compared with a 5.8% mean increase in score for patients treated with placebo ($p < 0.001$) at Week 12. A mean reduction of 57.2% was recorded for the 23 DVU score in patients treated with IFX compared to a mean reduction of 3.4% in patients given placebo ($p < 0.001$). Readers agreed that 6/16 (37.5%) patients who received IFX still had residual inflammation on MRI when all 23 spinal units were assessed.

Guyatt's effect size for the change in SPARCC MRI spinal inflammation scores at Week 12 was well over 1 for both the 6 DVU and 23 DVU scores by each of the 3 read-

Table 1. Patient demographic and baseline characteristics.

Baseline Characteristic	Placebo, n = 18	Infliximab, n = 18
Age, yrs, mean (SD)	41.7 (9.3)	43.6 (11.8)
Male, %	77.8	77.8
Years since first symptoms, mean (SD)	20.2 (11.5)	20.0 (11.8)
Years since first diagnosis, mean (SD)	14.3 (12.0)	12.0 (11.2)
HLA B27-positive, %	72.2	77.8
History of inflammatory bowel disease, %	11.1	11.1
History of uveitis, %	33.3	22.2
History of psoriasis, %	5.6	5.6
BASDAI, mean (SD)	6.8 (1.4)	6.6 (1.2)
BASFI, mean (SD)	6.1 (2.0)	6.9 (1.4)
BASGI, mean (SD)	6.7 (1.5)	7.4 (1.0)
ESR, mm/h, mean (SD)	31.6 (21.8)	21.1 (12.9)
CRP, mg/l, mean (SD)	25.7 (23.8)	15.3 (13.7)

BASDAI: Bath Ankylosing Spondylitis Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; BASGI: Bath Ankylosing Spondylitis Global Index; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein.

Table 2. SPARCC MRI 6 DVU scores for spinal inflammation at baseline and at 12 weeks after randomization to either placebo (n = 16) or infliximab 3 mg/kg (n = 16) in patients with AS.

Reader	Treatment	Baseline mean (SD) median (25/75 IQ) 95% CI range	Week 12 mean (SD) median (25/75 IQ) 95% CI range	Change mean (SD) median (25/75 IQ) 95% CI range	p
A	Placebo	19.00 (16.77)	18.88 (19.01)	-0.13 (6.98)	0.0012
		16.50 (4.25/34.75)	13.00 (1.25/35.50)	0.0 (-2.75/1.0)	
		10.06, 27.94	8.75, 29.00	-3.84, 3.59	
A	Infliximab	0.0-47.0	0.0-58.0	-18.0 to 15.0	
		18.94 (17.98)	7.0 (9.89)	-11.94 (12.24)	
		13.5 (1.25/41.75)	1.50 (0.0/10.0)	-9.50 (-16.0/-1.0)	
B	Placebo	9.36, 28.52	1.73, 12.27	-18.46, -5.42	0.0009
		0.0-46.0	0.0-32.0	-39.0 to 0.0	
		17.19 (21.43)	16.69 (21.29)	-0.50 (6.55)	
B	Infliximab	5.50 (2.00/40.00)	6.50 (2.00/38.00)	0.0 (0.0-0.75)	
		5.77, 28.61	5.34, 28.03	-3.99, 2.99	
		0.0-61.00	0.0-62.0	-23 to 10.0	
C	Placebo	14.13 (16.89)	3.94 (8.42)	-10.19 (12.07)	0.0014
		9.00 (0.00/24.00)	0.00 (0.0/6.00)	-7.00 (-17.00/0.00)	
		5.12, 23.13	-0.55, 8.42	-16.62, -3.75	
C	Infliximab	0.00-51.00	0.00-33.00	-36.00 to 0.00	
		17.06 (16.04)	17.19 (17.84)	0.13 (5.06)	
		12.0 (4.25/26.75)	9.00 (3.25/27.0)	0.0 (-2.75/1.75)	
C	Infliximab	8.52, 25.61	7.68, 26.69	-2.57, 2.82	
		0.0-52.0	0.0-57.0	-10 to 11.0	
		16.31 (16.60)	5.44 (9.47)	-10.88 (9.32)	
C	Infliximab	11.50 (1.75/23.0)	1.0 (0.0/5.5)	-10.0 (-17.50/-0.25)	
		7.47, 25.16	0.39, 10.49	-15.84, -5.91	
		0.0-50.0	0.0-29.0	-31 to 0.0	

SPARCC: Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; DVU: disc-vertebral units; AS: ankylosing spondylitis. IQ: interquartile.

Table 3. SPARCC MRI 23 DVU scores for spinal inflammation at baseline and at 12 weeks after randomization to either placebo (n = 16) or infliximab 3 mg/kg (n = 16) in patients with AS.

Reader	Treatment	Baseline mean (SD) median (25/75 IQ) 95% CI range	Week 12 mean (SD) median (25/75 IQ) 95% CI range	Change mean (SD) median (25/75 IQ) 95% CI range	p
B	Placebo	24.13 (33.83)	23.56 (34.11)	-0.56 (6.51)	0.0009
		5.5 (2.0/55.0)	6.5 (2.0/50.75)	0.0 (0.0/0.75)	
		6.10, 42.15	5.39, 41.74	-4.03, 2.91	
B	Infliximab	0.0-94.0	0.0-97.0	-23 to 10	
		16.56 (21.98)	4.56 (9.71)	-12.0 (16.26)	
		9.0 (0.0/24.0)	0.0 (0.0/6.0)	-7.0 (-18.5/0.0)	
C	Placebo	4.85, 28.27	-0.61, 9.74	-20.67, -3.33	0.0016
		0.0-70.0	0.0-37.0	-54.0 to 0.0	
		23.88 (26.85)	24.44 (30.43)	0.56 (5.81)	
C	Infliximab	12.50 (4.25/41.75)	9.0 (3.25/39.0)	-0.50 (-2.75/3.25)	
		9.57, 38.18	8.22, 40.65	-2.53, 3.66	
		0.0-90.0	0.0-94.0	-10.0 to 15	
C	Infliximab	22.50 (26.09)	7.69 (13.05)	-14.81 (14.74)	
		11.50 (1.75/35.5)	1.50 (0.0/9.25)	-10.0 (-27.25/-0.25)	
		8.60, 36.40	0.74, 14.64	-22.67, -6.96	
C	Infliximab	0.0-77.0	0.0-39.0	-50.0 to 0.0	

SPARCC: Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; DVU: disc-vertebral units; AS: ankylosing spondylitis. IQ: interquartile.

ers indicating a large treatment effect by IFX in reducing spinal inflammation (Table 4). A substantially greater effect size was noted for the 23 DVU method by 1 reader, although the second reader recorded the converse.

Reliability of SPARCC 6 DVU and 23 DVU scoring methods. The SPARCC 6 DVU and 23 DVU scoring systems demonstrated excellent and comparable reproducibility (Table 5). In particular, interreader ICC for the change in spinal inflammation from baseline to 12 weeks was 0.89 and 0.93 for the 6 DVU and 23 DVU scores, respectively, when recorded by the same 2 readers.

Association between SPARCC scores and clinical data. A reduction of acute-phase reactants from baseline to 12 weeks was significantly associated with change in spinal inflammation recorded by MRI, particularly when the 6 DVU method was used (Table 6). No significant association was noted when change in other clinical measurements was analyzed, in particular the BASDAI.

DISCUSSION

We show that low-dose IFX therapy is highly effective in reducing spinal inflammation observed on MRI. This effect was recorded in all patients who had inflammation at baseline. We also show that methods for quantifying spinal inflammation by MRI are equally and highly discriminatory, whether assessment is limited to only the most severely affected regions of the spine (6 DVU score) or includes the entire spine (23 DVU score). The data also reflect the dis-

cordance between MRI-defined and clinically defined outcome measures, which has been observed^{4,19}.

The degree of reduction in spinal inflammation using the SPARCC method is comparable to that reported for adalimumab at 12 weeks, with reduction from baseline scores of 54% compared to an increase in the placebo group of 9% using the SPARCC 6 DVU method¹⁹. This is similar to reports assessing patients on higher doses of IFX (5 mg/kg) but using a different scoring method and assessing patients at 24 weeks²⁰. A decrease in spinal inflammation has also been reported with etanercept after 24 weeks, but the largest study of etanercept included MRI of only the lumbar spine and lower thoracic spine and MRI was limited to only 40 patients²¹. Open-label extension of these pivotal trials has shown that the reduction in spinal inflammation is persistent^{19,20}.

We have reported the lack of association with clinical measurements of disease activity but have consistently shown significant correlations with acute-phase reactants^{4,19,22}. In particular, change in MRI acute lesions correlated significantly with change in CRP in patients receiving adalimumab among 82 patients recruited to a placebo-controlled trial¹⁹. A report that examined the effects of IFX in a 20-patient subgroup analysis of the ASSERT trial described a correlation between change in BASDAI score and change in acute lesion score as measured by the ASspiMRI method³, but did not report correlations with acute-phase reactants. A second report of 20 patients that formed a subgroup from another controlled trial of IFX demonstrated only borderline significance for the correlation with the BASDAI after 12 weeks of treatment and no correlation with either clinical or laboratory measurements after 2 years of followup²³. Similarly, improvement of the MRI scores did not show a significant correlation with changes in clinical or laboratory measurements in 2 controlled trials of etanercept, although the analyses were limited to a subgroup of only 40 patients in the first study²¹ and only 15 patients in the second²⁴. In addition, the MR scan of the spine was limited to the lower thoracic and lumbar regions in the former study²¹.

Our data therefore suggest that symptoms are unrelated to the presence of bone marrow edema in the vertebral bodies, which is the primary feature assessed by MRI in currently used scoring methods. Other sources of symptoms may relate to inflammation in other structures such as the posterior elements and entheses, or noninflammatory causes of pain, such as secondary mechanical factors. These data also show that MRI is highly sensitive in the detection of inflammation, and the lack of sensitivity of acute-phase reactants and their relatively modest correlation with MRI indicates that they cannot be substituted for MRI examination. This in turn brings into question the clinical significance of MRI and whether the lesions observed on MRI reflect disease process and clinical outcomes. This is high-

Table 4. Guyatt's effect size per reader per method. Guyatt's effect size is the mean difference in the treatment group divided by the standard deviation of the change in the placebo group. Effect size > 0.8 is considered large.

Method	Reader	Guyatt's Effect Size
6 DVU	A	1.7
6 DVU	B	2.0
6 DVU	C	1.7
23 DVU	B	1.8
23 DVU	C	2.6

DVU: discovertebral units.

Table 5. Interreader reliability of status and change in SPARCC MRI spinal inflammation scores.

Method	Pretreatment Status ICC	Posttreatment Status ICC	Change ICC
Readers A and C 6 DVU	0.89	0.95	0.81
Readers A and B 6 DVU	0.89	0.94	0.91
Readers B and C 6 DVU	0.94	0.95	0.89
Readers B and C 23 DVU	0.92	0.92	0.93

ICC: intraclass correlation coefficient; SPARCC: Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; DVU: discovertebral units.

Table 6. Correlation matrix for change in clinical and SPARCC MRI spinal inflammation scores.

SPARCC MRI Score	Δ BASDAI	Δ BASFI	Change in Δ ESR	Change in Δ CRP
Δ 6 DVU*	0.25	0.16	0.57 (p = 0.001)	0.45 (p = 0.012)
Δ 6 DVU**	0.14	0.07	0.43 (p = 0.02)	0.34
Δ 23 DVU**	0.17	0.14	0.44 (p = 0.017)	0.31

* Mean of reader A and C scores. ** Mean of reader B and C scores. SPARCC: Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; BASDAI: Bath Ankylosing Spondylitis Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; DVU: discovertebral units.

lighted in the comparison of the MRI changes in the responder and nonresponder groups in our study. Limited histopathological data have shown that lesions observed on STIR sequences correlate modestly with histopathological grades of inflammation²⁵. Moreover, the presence of increased STIR signal at vertebral corners is associated with the subsequent development of syndesmophytes on radiography after 2 years of followup²⁶. Consequently, assessment of spinal inflammation using MRI should continue to be a requirement for the development of new therapeutics for AS.

Our comparative analysis of the 6 DVU and 23 DVU methods has also demonstrated equally high interobserver reliability and the ability to discriminate between treatment groups. This consistency in reliability and responsiveness between readers has been reported⁴, and is particularly noteworthy because this was observed in readers with no prior experience with any MRI-based scoring system². The high reliability may be attributed to the dichotomous scheme for recording bone marrow edema, a clear definition of what constitutes abnormal STIR signal, and the availability of reference images that visually outline the measurements of an abnormal STIR signal. This clearly outweighs any potential for diminished reliability that might be anticipated by the requirement to select the 6 most severely affected DVU. Conversely, the requirement to score all 23 DVU mandates the scoring of regions that may be questionable with respect to size and presence of increased STIR signal⁵. In addition, certain MRI artifacts may resemble increased STIR signal. Phase-encoding artifacts due to blood flow in the major vessels of the abdomen is a common source of misinterpretation with inflammation in the anterior portion of the lumbar vertebral bodies. The high responsiveness and ability to discriminate between treatment groups reflects both the efficacy of treatment on local inflammation and the approach taken by the SPARCC method to record the extent of the lesion in 3 dimensions by requiring assessment of the lesion in consecutive sagittal slices. This capability is unique to MRI among imaging modalities. Moreover, the SPARCC method also applies an added weighting for the depth and intensity of the lesion. The minimal difference between the 6 DVU and 23 DVU methods may reflect the observation that scoring 6 DVU identifies all the detectable inflamma-

tion in the anterior spine in two-thirds of patients. The smaller lesions that might be present in the remaining DVU are likely to demonstrate much less relative change following treatment. Since scoring all 23 DVU is time-consuming, particularly when evaluating questionable lesions, we continue to recommend the scoring of only the most severely affected spinal regions when using the SPARCC MRI method to assess therapeutics in clinical trials. The assessment of all 23 DVU may be more appropriate in observational cohort studies, where questions beyond therapeutic efficacy are being addressed. These data indicate that the SPARCC 23 DVU method can be used reliably for observational studies and demonstrates a high degree of responsiveness.

We have shown that low-dose IFX is highly effective in reducing spinal inflammation recorded by MRI. Inflammation does not need to be scored in the entire spine using MRI. Discrimination between treatment groups and reliability between observers is equally high when scoring focuses on only the most severely affected regions of the spine. Feasibility is also improved by the use of appropriate viewing conditions and the simultaneous availability of electronic data entry on user-friendly schematic Web pages designed to depict the anatomical region scored on MRI.

REFERENCES

1. Rennie WJ, Dhillon SS, Conner-Spady B, Maksymowych WP, Lambert RGW. MRI assessment of spinal inflammation in ankylosing spondylitis: standard clinical protocols may omit inflammatory lesions in thoracic vertebrae. *Arthritis Rheum* 2009;61:1187-93.
2. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Ostergaard M, et al. ASAS/OMERACT MRI in AS Working Group. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. *J Rheumatol* 2007;34:862-70.
3. Braun J, Baraliakos X, Golder W, Brandt J, Rudwaleit M, Listing J, et al. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab: evaluation of a new scoring system. *Arthritis Rheum* 2003;48:1126-36.
4. Maksymowych WP, Dhillon SS, Park R, Salonen D, Inman RD, Lambert RGW. Validation of the Spondyloarthritis Research Consortium of Canada (SPARCC) MRI Spinal Inflammation Index: Is it necessary to score the entire spine? *Arthritis Rheum* 2007;57:501-7.

5. Maksymowych WP, Lambert RGW. Magnetic resonance imaging for spondyloarthritis — Avoiding the minefield [editorial]. *J Rheumatol* 2007;34:259-65.
6. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Krishnananthan R, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. *Arthritis Rheum* 2005;53:502-9.
7. Baraliakos X, Landewe R, Hermann KG, Listing J, Golder W, Brandt J, et al. Inflammation in ankylosing spondylitis: a systematic description of the extent and frequency of acute spinal changes using magnetic resonance imaging. *Ann Rheum Dis* 2005;64:730-4.
8. Braun J, Brandt J, Listing J, Zink A, Alten R, Golder W, et al. Treatment of active ankylosing spondylitis with infliximab: a randomized controlled multicentre trial. *Lancet* 2002;359:1187-93.
9. van der Heijde D, Dijkmans B, Geusens P, Sieper J, DeWoody K, Williamson P, et al. Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). *Arthritis Rheum* 2005;52:582-91.
10. Maksymowych WP, Jhangri GS, Lambert RG, Mallon C, Buenaviaje H, Pedrycz E, et al. Infliximab in ankylosing spondylitis: an inception prospective observational cohort analysis of efficacy and safety. *J Rheumatol* 2002;29:959-67.
11. Keeling S, Oswald A, Russell AS, Maksymowych WP. Prospective observational analysis of the efficacy and safety of low-dose (3 mg/kg) infliximab in ankylosing spondylitis: 4-year followup. *J Rheumatol* 2006;33:558-61.
12. Inman RD, Maksymowych W, for the CANDLE Study Group. A double-blind, placebo-controlled trial of low dose infliximab in ankylosing spondylitis. *J Rheumatol* 2010;37:1203-10.
13. 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.
14. Garrett S, Jenkinson T, Kennedy G, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *J Rheumatol* 1994;21:2286-91.
15. Anderson JJ, Baron G, van der Heijde D, Felson DT, Dougados M. Ankylosing Spondylitis Assessment Group preliminary definition of short-term improvement in ankylosing spondylitis. *Arthritis Rheum* 2001;44:1876-86.
16. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. *J Rheumatol* 1994;21:2281-5.
17. van der Heijde D, Dougados M, Davis J, Weisman MH, Maksymowych W, Braun J, et al. ASsessment in Ankylosing Spondylitis International Working Group/Spondylitis Association of America recommendations for conducting clinical trials in ankylosing spondylitis. *Arthritis Rheum* 2005;52:386-94.
18. Sloan J, Symonds T, Vargas-Chanes D, Fridley B. Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. *Drug Inf J* 2003;37:23-31.
19. Lambert RGW, Salonen D, Rahman P, Inman RD, Wong RL, Einstein SG, et al. Adalimumab significantly reduces both spinal and sacroiliac joint inflammation in patients with ankylosing spondylitis. *Arthritis Rheum* 2007;56:4005-14.
20. Braun J, Landewé R, Hermann KG, Han J, Yan S, Williamson P, et al. Major reduction in spinal inflammation in patients with ankylosing spondylitis after treatment with infliximab: results of a multicenter, randomized, double-blind, placebo-controlled magnetic resonance imaging study. *Arthritis Rheum* 2006;54:1646-52.
21. Baraliakos X, Davis J, Tsuji W, Braun J. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis before and after therapy with the tumor necrosis factor alpha receptor fusion protein etanercept. *Arthritis Rheum* 2005;52:1216-23.
22. Maksymowych WP, Dhillon SS, Park R, Salonen D, Inman RD, Lambert RGW. Validation of the Spondyloarthritis Research Consortium of Canada (SPARCC) MRI Spinal Inflammation Index: Is it necessary to score the entire spine? *Arthritis Rheum* 2007;57:501-7.
23. Sieper J, Baraliakos X, Listing J, Brandt J, Haibel H, Rudwaleit M, et al. Persistent reduction of spinal inflammation as assessed by magnetic resonance imaging in patients with ankylosing spondylitis after 2 yrs of treatment with the anti-tumour necrosis factor agent infliximab. *Rheumatology* 2005;44:1525-30.
24. Baraliakos X, Brandt J, Listing J, Haibel H, Sorensen H, Rudwaleit M, et al. Outcome of patients with active ankylosing spondylitis after two years of therapy with etanercept: clinical and magnetic resonance imaging data. *Arthritis Rheum* 2005;53:856-63.
25. Appel H, Loddenkemper C, Grozdanovic Z, Ebhardt H, Dreimann M, Hempfing A, et al. Correlation of histopathological findings and magnetic resonance imaging in the spine of patients with ankylosing spondylitis. *Arthritis Res Ther* 2006;8:R143.
26. Maksymowych WP, Chiowchanwisawakit P, Clare T, Pedersen SJ, Østergaard 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 ankylosis. *Arthritis Rheum* 2008;60:93-102.