

Magnetic Resonance Imaging Assessment of Spinal Inflammation in Patients Treated for Ankylosing Spondylitis

MARCUS TREITL, MARKUS KORNER, CHRISTA BECKER-GAAB, MALTE TRYZNA, JOHANNES RIEGER, KLAUS-JUERGEN PFEIFER, MAXIMILIAN F. REISER, and STEFAN WIRTH

ABSTRACT. Objective. To compare different magnetic resonance imaging (MRI) based algorithms for assessment of spinal inflammation in patients with ankylosing spondylitis (AS) being treated with disease modifying drugs.

Methods. Eleven patients (10 men, 1 woman) who fulfilled modified New York diagnostic criteria and had severe disease [Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) > 4] were given intravenous infusion of infliximab (Remicade[®], 5 mg/kg) for 96 weeks. Whole-spine MRI was done at 0, 24, and 54 weeks. Measurements of the Ankylosing Spondylitis Spinal MRI Activity Score (ASspiMRI), paravertebral inflammatory lesion count (pILC), contrast:noise ratio (CNR) measurements of defined inflammatory lesions, and other scores together with C-reactive protein concentration were made at each visit. Examinations were anonymized and randomly presented twice to 2 radiologists. The significance of any changes in scores, their correlation with the BASDAI, and interobserver and intraobserver correlations were calculated.

Results. The mean (\pm SD) BASDAI improved from 7.2 (\pm 1.5) to 1.3 (\pm 0.9) after 54 weeks ($p < 0.001$), and the ASspiMRI score improved from 12.0 (\pm 8.0) to 0.2 (\pm 0.5) ($p < 0.001$). Correlations between ASspiMRI score and BASDAI were 0.831, 0.746, and 0.369 ($p < 0.001$ each). The pILC improved significantly ($p < 0.01$). CNR showed no correlation with any clinical score.

Conclusion. The ASspiMRI score performed best for assessment and quantification of spinal inflammation and disease activity in patients with AS, but should also quantify paravertebral inflammatory lesions, since we could show that this will significantly improve its correlation to clinical scores and increase its sensitivity to mild inflammatory processes. (First Release Dec 1 2007; J Rheumatol 2008;35:126–36)

Key Indexing Terms:

ANKYLOSING SPONDYLITIS
SPINE

MAGNETIC RESONANCE IMAGING
SPONDYLOARTHROPATHY

INFLIXIMAB

Ankylosing spondylitis (AS) is a chronic rheumatic disease characterized by axial skeletal ankylosis, enthesitis^{1,2}, peripheral arthritis, and expression of the HLA-B27 molecule^{3,4}. It is one of the most common rheumatic diseases and the main constituent of the spondyloarthritides. The disease predominantly affects young men during the third decade⁵; the male:female ratio is 3:1. The key clinical features are inflammatory back pain, sacroiliitis^{6,7}, and spondylitis, in the course of which inflammation progresses to formation of new bone, and ankylosis is common^{6,8}. For a long time it was considered to be a relatively benign form of arthritis⁹, but we now know that it causes degrees of pain and disability similar to rheuma-

toid arthritis¹⁰, and that many patients have severe inflammatory symptoms even decades after diagnosis¹¹⁻¹³. There is therefore a great demand for an effective antiinflammatory treatment that will prevent the progression of spinal lesions and ankylosis¹⁴, and for imaging techniques that will identify and quantify both the state of disease and the therapeutic effects early and reliably¹⁵.

New disease modifying drugs such as infliximab (Remicade[®]) and etanercept (Enbrel[®]), both of which act by differential blockade of tumor necrosis factor- α (TNF- α), have the potential to alter the course of the disease appreciably; several clinical trials have shown that their antiinflammatory effect is considerable¹⁶⁻¹⁹ and they are able to reduce the formation of new bone and the risk of ankylosis^{20,21}. For the first time we have drugs that can slow the progression of axial disease in these patients¹³.

However, we still need a technique that will quantify the course of morphological and inflammatory changes in bones, joints, and soft tissue to objectively assess the effects of these drugs²². For decades, plain radiographs were used to detect inflammatory and bony lesions²³, but attempts to introduce a

From the Department for Clinical Radiology, Clinical Centre of Ludwig-Maximilians-University of Munich, Munich, Germany.

M. Treitl, MD; M. Korner, MD; C. Becker-Gaab, MD; M. Tryzna MD; J. Rieger, MD; K.-J. Pfeifer, MD, PhD; M. Reiser, MD, PhD; S. Wirth, MD.

Address reprint requests to Dr. M. Treitl, Department for Clinical Radiology, University of Munich, Pettenkoferstrasse 8a, D-80336 Munich, Germany. E-mail: Marcus.Treitl@med.uni-muenchen.de

Accepted for publication August 22, 2007.

reliable score for the quantification of improvements failed because of the lack of sensitivity to change^{24,25}.

In contrast, native and contrast-enhanced magnetic resonance imaging (MRI) of the spine permits not only documentation of late stages of the disease, but also depiction of early inflammatory lesions, because it gives excellent soft-tissue contrast²⁶⁻²⁸. It would therefore be an ideal tool for the morphological and quantitative assessment of improvements²⁹⁻³¹.

Braun, *et al* recently developed the Ankylosing Spondylitis Spinal MRI Activity and Chronicity Score (ASspiMRI) for quantitative assessment of acute and chronic changes in the spine of patients with AS^{15,22}, which is good for assessing acute inflammation, but gives only acceptable correlation with clinical improvement as assessed by the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)^{15,32,33}. It assesses changes in vertebral bodies but does not detect paravertebral inflammatory lesions. We know from experience that some patients may have only paravertebral inflammation, with none in the vertebral body itself, so this will lead to invalidation of the score, and probably explains the unexpectedly mediocre performance of this instrument. MRI of the spine offers not only the possibility of depicting structural changes but also of quantifying contrast enhancement and calculating alterations in signal. These features are also not covered by the ASspiMRI.

MRI may play a critical part in fast and early assessment of structural and inflammatory changes at the spine in these patients, so there is an increasing demand for a reliable, validated score. We evaluated and compared several contemporary algorithms and scores for MRI-based assessment of spinal inflammation in patients with AS, particularly the ASspiMRI activity score, to show possible improvement by considering paravertebral inflammatory lesions and measurements of signal:noise ratio (SNR) and contrast:noise ratio (CNR).

MATERIALS AND METHODS

Eleven patients (10 male, 1 female) who fulfilled the modified New York classification criteria for AS²³ and had active disease as defined by a BASDAI³⁴ score > 4 and spinal pain \geq 4 recorded on a visual analog scale (VAS) were studied. The mean age was 37 years (range 24–50) (Table 1). Because results

of the evaluation of the ASspiMRI activity score had to be comparable with those of other publications that used this score, the study protocol was adopted from the first publication of the ASspiMRI activity score by Braun, *et al*¹⁵.

Randomization was double-blind, and patients were allocated to one of 2 groups. The first group received a placebo infusion at Weeks 0, 2, 6, 12, and 18; then they received an infusion of infliximab 5 mg/kg intravenously (Remicade®; Centocor Inc., Horsham, PA, USA) at Weeks 24, 26, 30, 36, 42, 48, and 54 until Week 96. The second group received an infliximab infusion 5 mg/kg intravenously at weeks 0, 2, 6, 12, 18, 24, 26, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84, 90, and 96. Because the first (placebo) group received infliximab after 18 weeks, and since only 2 patients were in this group, we evaluated all patients as one group. Patients had clinical examination and MRI before the first infusion of infliximab and at Weeks 24, 54, and 102. At each visit the BASDAI, Bath AS Metrology Index (BASMI) and Bath AS Functional Index (BASFI), C-reactive protein (CRP) concentrations, and VAS for night pain were recorded by an experienced rheumatologist. In addition the Maastricht Ankylosing Spondylitis Enthesitis Score (MASES) was applied. The mean score was 22.1 at Week 0. Unfortunately, no followup data for this score were available. At Week 0 plain radiographs of the whole spine and the sacroiliac joints were screened for radiographic features of AS by 2 experienced radiologists by consensus (Table 1). Eight of the patients had signs of sacroiliitis, with a mean degree of 2.61 according to the New York scoring method for the sacroiliac joints (0 = no abnormalities; 1 = suspicious changes but no specific abnormalities; 2 = minimal sacroiliitis, loss of definition at the edge of the sacroiliac joints, some sclerosis and perhaps minimal erosions, and perhaps some joint space narrowing; 3 = moderate sacroiliitis, definite sclerosis on both sides of the joint, blurring and indistinct margins, and erosive changes with loss of joint space; 4 = complete fusion or ankylosis of the sacroiliac joint without any residual sclerosis; Table 1). Ankylosis of the spine or the sacroiliac joints was found in only one (9.1%; Table 1).

Written informed consent for participation in the study was obtained from all patients. The study was approved by the local research ethics committee of the university hospital, and complied with the Declaration of Helsinki.

MRI of the spine. MRI was taken at 1.5 Tesla (Magnetom Vision, Siemens Medical Solutions, Erlangen, Germany) with a dedicated spine coil. To show the whole spine, 2 distinct examinations were required with a standard interval of 2 days. The cervical and thoracic spines were examined at the first session, the lumbar spine at the second. For each region 3 pulsed sequences were obtained in the sagittal plane: T1-weighted before and T1-weighted fat-saturated after Gd-DTPA (0.1 mmol/kg; Schering, Berlin, Germany) and a short-tau inversion recovery (STIR) sequence (Table 2). To simplify comparison of examinations of different timepoints, high-intense MRI markers (MR-SPOTS®; Beekley Corp., Bristol, CT, USA) were attached to patients' skin at the position of the second cervical, the eighth and tenth thoracic, and the first sacral vertebral body. Because ASspiMRI activity scores had already reached zero values at Week 54, only examinations at Weeks 0, 24, and 54 were included in the results.

Table 1. Baseline characteristics of the study population (n = 11).

Median age (range), yrs	37 (24–50)
Male:female, n	10:1
Median duration of disease (range), yrs	10.3 (1.6–18.9)
No. HLA-B27-positive*	10
Mean CRP level at screening (range), IU/ml	29.6 (8–99)
No. patients with involvement of peripheral joints	1
MASES score	22.2 (2–36)
No. patients with radiographic sacroiliitis	8
Mean degree of sacroiliitis (determined radiographically, range)	2.61 (1–4)
No. patients with ankylosis of the spine, or sacroiliac joints, or both	1
Mean BASDAI at screening (range)	6.8 (4.3–8.9)
Mean weight at screening (range) kg	84.6 (59–100)

* Test positive if titer \geq 100 IU/ml. MASES: Maastricht Ankylosing Spondylitis Enthesitis Score; BASDAI: Bath AS Disease Activity Index.

Table 2. Characteristics of pulse sequences used for spinal MRI. MRI of the upper and lower part of the spine were done within 2 days using the same scanner. Mean duration of examinations was 20 minutes (range 18–25).

Sequence	Upper Spine (C2–T10)			Lower Spine (T8–S1)		
	TSE T1w without Gd-DTPA	STIR	TSE T1w fat sat with Gd-DTPA	TSE T1w without Gd-DTPA	STIR	TSE T1w fat sat with Gd-DTPA
Plane	Sagittal	Sagittal	Sagittal	Sagittal	Sagittal	Sagittal
Field of view ⁴¹	34–38	34–38	34–38	34–38	34–38	34–38
Matrix size, pixels	512×256	512×256	512×256	512×256	512×256	512×256
TE, ms	15	45	15	15	45	15
TR, ms	650	4000	650	650	4000	650
TI, ms	—	150	—	—	150	—
Slice thickness, mm	3/0	3/0	3/0	3/0	3/0	3/0

TSE: turbo spin-echo; STIR: short-tau inversion recovery; T1w: T1-weighted; Fat sat: fat saturated.

Reading the radiographs. After completion of the study the MR images were presented to 2 experienced radiologists. Images were presented randomly, and the readers were unaware of patients' names, date of examination, clinical history, and each others' opinions. To assess the intraobserver correlation each radiologist had to read all MRI examinations twice. There was an interval of 8 weeks between the 2 sessions.

MRI scoring with ASspiMRI. MR images were analyzed using the ASspiMRI activity score^{15,22,23}. The ASspiMRI chronicity score was abandoned because it assessed only irreversible lesions that could not be altered by a disease modifying drug, and needed longer observation periods to detect even minor changes. In addition only one patient of the study group had ankylosis on plain radiographs (Table 1). Disease activity was assessed by contrast enhancement in T1-weighted images after injection of Gd-DTPA, and bone edema in the corresponding STIR sequence. For application of the ASspiMRI activity score the spine was divided into 23 vertebral units between the middle of the second cervical and the middle of the first sacral vertebral body. One unit was defined as the region between 2 virtual lines through the middle of each vertebral body^{15,22,23}. The maximum possible ASspiMRI activity score was 138, as each vertebral unit was graded as follows: 0 = normal; 1 = enhancement/edema ≤ 25%; 2 = enhancement/edema ≤ 50%; 3 = enhancement/edema ≥ 50%; 4 = small erosion ≤ 25%; 5 = erosion > 25% < 50%; and 6 = large erosion ≥ 50%.

Assessment of paravertebral lesions. Because the ASspiMRI activity score grades only lesions in the vertebral body itself, we examined 3 additional, anatomically exactly definable, paravertebral regions to assess paravertebral inflammatory processes. The regional allocation was based on 10 years' experience in evaluation of spinal MRI in patients with AS. The following paravertebral units were defined: spinal processes of C2 to S1 (SP), small intervertebral joints of C2 to S1 (SIJ), and costovertebral joints of T1 to T12 (CVJ). Each unit comprises a strip 1 cm wide of surrounding soft tissue, and the presence (= 1) or absence (= 0) of inflammatory processes was assessed and the results of all sites were summed, giving a maximum score of 116 for the total count. In addition, simple modifications were made to the ASspiMRI activity score by inclusion of the numbers of affected paravertebral units.

Assessment of changes by enhancement with contrast media. To evaluate the influence of changes in the SNR and CNR of defined inflammatory lesions on clinical improvements, we defined a representative contrast-enhanced vertebral or paravertebral lesion and evaluated it at followup. We measured the signal intensity within this lesion and a reference lesion in unaffected bone marrow in T1-weighted fat-saturated contrast-enhanced images using a standard and reproducible region of interest with the standard tools of the picture archive and communication system (PACS; Impax R4.1, Agfa Gevaert GmbH, Cologne, Germany). The region of interest was set in the slice with the greatest extent of the lesion. The values for signal intensity and noise of the representative and reference lesions were used to calculate the SNR and CNR by the following formula:

$$\text{SNR} = \text{signal intensity of lesion/noise}$$

$$\text{CNR} = (\text{signal intensity of lesion} - \text{signal intensity of bone marrow})/\text{noise}$$

Statistical analysis. For statistical analyses SPSS 15.0[®] was used (SPSS, Chicago, IL, USA). For each timepoint (weeks 0, 24, and 54) the difference of BASDAI, BASFI, BASMI, CRP, night pain VAS, ASspiMRI, paravertebral inflammatory lesion count, SNR, and CNR was calculated, and the statistical significance of differences determined using Wilcoxon's signed-rank test for 2 paired samples (analysis of improvement over time) and Friedman's 2-way analysis of variance for paired samples (for all 3 timepoints at once). Changes in ASspiMRI activity score, paravertebral inflammatory lesion count, SNR, and CNR were correlated with BASDAI as an indicator of clinical improvement and CRP concentrations as an indicator of inflammatory reaction by multiple linear regression analyses and calculation of Pearson's correlation coefficient *r*. The statistical 2-sided significance of the calculated correlation is expressed as *p* values at a level of significance of 0.01. Interobserver and intraobserver correlation was assessed by intraclass correlation analyses for all 3 timepoints. In addition, intraobserver correlation was evaluated graphically by Bland-Altman plots. The mean values of the results of both readers are graphed against the difference, and the mean value and 2-fold standard deviation are marked. Scattering of data points around the mean value correlates with the broadness of interobserver variability.

RESULTS

Clinical scores after 54 weeks of intermittent intravenous infliximab therapy. All patients completed the study protocol until Week 54, including all followup examinations. The mean BASDAI score at Week 0 was 7.2 (± 1.5) (Table 3, Figure 1). There was a significant decrease in the mean BASDAI score to 1.3 (± 0.9) until Week 54 (*p* < 0.001). The reduction in BASDAI score between Weeks 24 and 0 was significant (*p* < 0.001) as was that between Weeks 54 and 0 (*p* < 0.001), but not that between Weeks 54 and 24 (*p* < 0.065; Table 3). There was a slight delay in clinical improvement for the 2 patients of the placebo group until Week 24 (data not shown), which was equalized by Week 54.

Values for the BASFI score decreased from 5.6 (± 1.1) at the start to 1.8 (± 1.1; *p* < 0.001) until Week 54. Again the improvements over time were significant between Weeks 24 and 0 (*p* < 0.001) and between Weeks 54 and 0 (*p* < 0.001), but were less significant between Weeks 54 and 24 (*p* < 0.022). In contrast, there was only a minor and statistically

Table 3. Mean change and statistical significance of measured changes for clinical and MRI scores and for paravertebral units and signal:noise ratios (SNR) and contrast:noise ratios (CNR) of selected regions. Data for Week 0 do not include individuals from the placebo group (n = 2).

	Week 0		Week 24	p*	Mean ± SD	Week	Week 54	p*	Week	p*	Friedman
	Mean ± SD	Mean ± SD	Week		Mean ± SD	54-24, %	54-24, %		54-0, %		p**
Clinical scores											
BASDAI	7.2 ± 1.5	2.3 ± 1.7	68.1	< 0.001	1.3 ± 0.9	43.5	< 0.065	81.9	< 0.001	< 0.001	< 0.001
BASFI	5.6 ± 1.1	2.6 ± 1.6	53.6	< 0.001	1.8 ± 1.1	30.8	< 0.022	67.9	< 0.001	< 0.001	< 0.001
BASMI	4.2 ± 1.9	3.5 ± 2.0	16.7	< 0.063	3.5 ± 2.0	0.0	< 1.000	16.7	< 0.063	< 0.167	< 0.167
Night pain VAS	6.6 ± 2.9	2.7 ± 2.9	59.1	< 0.016	0.6 ± 0.7	77.8	< 0.008	90.9	< 0.001	< 0.001	< 0.001
CRP	29.5 ± 25.6	12.1 ± 31.7 [†]	59.0	< 0.021	0.5 ± 0.4	95.9	< 0.001	98.3	< 0.001	< 0.001	< 0.001
MRI score											
ASspiMRI activity score	12.0 ± 8.0	3.9 ± 4.8	67.5	< 0.001	0.2 ± 0.5	94.9	< 0.001	98.3	< 0.001	< 0.001	< 0.001
ROI:SNR/CNR											
SNR	9.9 ± 3.6	7.8 ± 3.4	21.2	< 0.001	5.5 ± 2.6	29.5	< 0.062	44.4	< 0.002	< 0.002	< 0.002
CNR	4.3 ± 2.4	1.3 ± 2.5	69.8	< 0.001	-0.8 ± 1.1	38.5	< 0.001	81.4	< 0.001	< 0.001	< 0.001
Paravertebral lesion counts											
Total for all regions	28.4 ± 2.8	7.3 ± 1.6	74.3	< 0.002	0.7 ± 0.3	90.4	< 0.016	97.5	< 0.001	< 0.001	< 0.001
Spinal processes	2.5 ± 2.3	0.5 ± 1.3	80.0	< 0.004	0.2 ± 0.6	60.0	< 0.25	92.0	< 0.002	< 0.001	< 0.001
Small intervertebral joints	1.7 ± 2.5	0.4 ± 0.9	76.5	< 0.031	0.0 ± 0.0	100.0	< 0.25	100.0	< 0.016	< 0.004	< 0.004
Costovertebral joints	2.9 ± 3.5	0.9 ± 2.4	69.0	< 0.008	0.0 ± 0.0	100.0	< 0.125	100.0	< 0.008	< 0.001	< 0.001

CRP: C-reactive protein. ROI: region of interest. * Exact significance, Wilcoxon signed-rank test. ** Exact significance, Friedman 2-way analysis of variance. † One patient presented unexpected high CRP levels at first followup due to acute influenza.

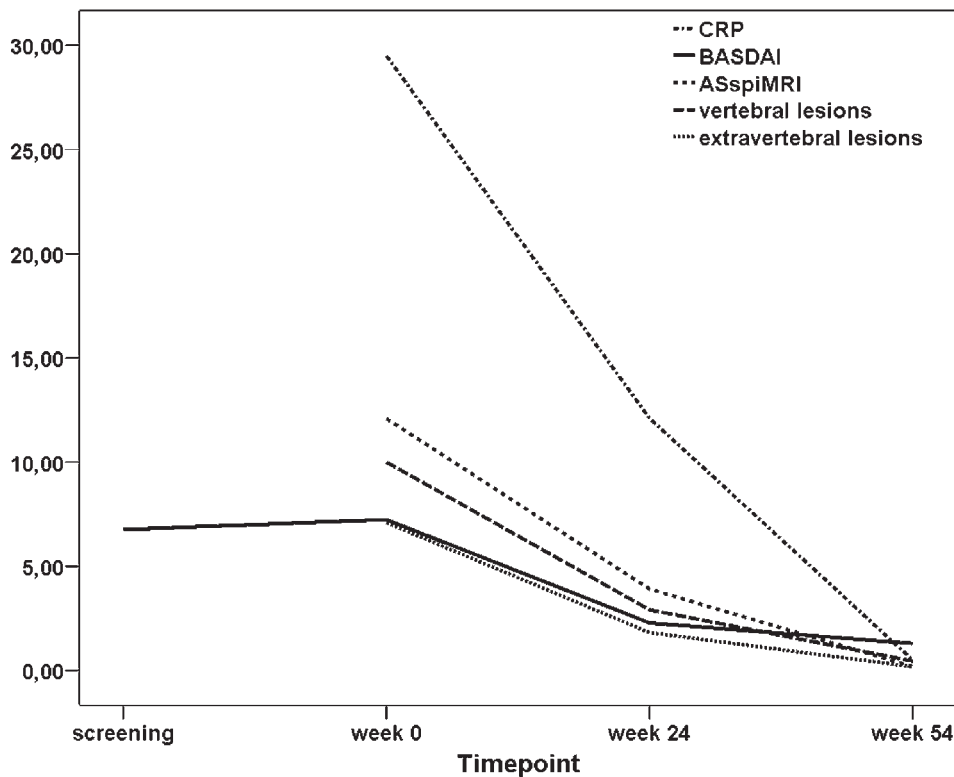


Figure 1. Time course of mean changes of clinical and MRI scores and CRP levels, and the count of vertebral and paravertebral lesions.

nonsignificant ($p < 0.167$) improvement of the BASMI from $4.2 (\pm 1.9)$ at study start to $3.5 (\pm 2.0)$ at Week 54 (Table 3).

A good response to infliximab was seen with night pain as assessed by VAS and with CRP levels. Both values did show a statistically significant decrease from $6.6 (\pm 2.9)$ and $29.5 (\pm 25.6)$ to $0.6 (\pm 0.7)$ and $0.5 (\pm 0.4)$, respectively ($p < 0.001$ each). For both these values the improvement over time was statistically significant at each timepoint (Table 3, Figure 1). Two of 11 patients (18.2%) had severe disease as indicated by high BASDAI levels > 8 , but did not exhibit any inflammatory lesion at the vertebral bodies. ASspiMRI activity score was ≤ 2 , consecutively (Figure 2). In these 2 cases, regions of inflammation were detected at the spinal processes only.

Evaluation of the ASspiMRI activity score

For 54 weeks of infliximab therapy and statistical evaluation. The ASspiMRI activity score was determined twice by both readers for all patients. Mean score at the start was $12.0 (\pm 8.0)$ and decreased to $0.2 (\pm 0.5)$ at Week 54. This shift was significant in total ($p < 0.001$) as well as for each improvement over time ($p < 0.001$ each; Table 3).

Correlation to BASDAI clinical score. The decrease of ASspiMRI activity score was correlated to clinical improvement as assessed by BASDAI score (Table 4, Figure 3). The correlation coefficient r of changes in ASspiMRI activity score and BASDAI scores was 0.831 between Weeks 24 and 0 ($p < 0.001$), 0.746 between Weeks 54 and 24 ($p < 0.001$), and 0.369 between Weeks 54 and 0 ($p < 0.001$).

Correlation to CRP levels. Values of the ASspiMRI activity score were correlated to CRP levels as well (Table 4, Figure 3). The correlation coefficient of changes in ASspiMRI activity score and CRP levels was 0.675 between Weeks 24 and 0 ($p < 0.023$), 0.414 between Weeks 54 and 24 ($p < 0.205$), and 0.636 between Weeks 54 and 0 ($p < 0.036$).

Interobserver and intraobserver correlation. Interobserver and intraobserver correlation was calculated for all data sets of the ASspiMRI activity score for both readers and each timepoint (Table 5). There were high values for the interobserver correlation at all timepoints (0.856, 0.924, and 0.751, respectively). Evaluation of interobserver correlation by Bland-Altman plots showed acceptable scattering around the mean value for all timepoints, and no outliers (Figure 4).



Figure 2. MRI of thoracic spine in the sagittal plane of a patient with inflammation exclusively in the paravertebral soft tissue at the spinal process (arrow), but with no inflammatory change at the vertebral bodies at Week 0. BASDAI was 8.9 and ASspiMRI activity score was 2. A. STIR sequence shows bone marrow edema at the spinal process. B. T1-weighted sequence before Gd-DTPA was given; C. T1-weighted sequence after Gd-DTPA, with strong contrast enhancement at the spinal process and the surrounding soft tissue as a result of acute inflammation (white arrows).

Table 4. Results of multiple linear regression analyses for ASspiMRI activity score, measurements of signal:noise ratios (SNR) and contrast:noise ratios (CNR) of selected lesions, and for count of paravertebral lesions to changes of BASDAI score and C-reactive protein (CRP) concentration. In addition, simple modifications of ASspiMRI activity score by addition of the paravertebral lesion counts were tested, resulting in a significant increase of correlation to clinical scoring and CRP levels. P values present 2-sided significance of r.

	Week 24-0			Week 54-24			Week 54-0		
	Correlation Coefficient, r	SE	p	Correlation Coefficient, r	SE	p	Correlation Coefficient, r	SE	p
Target correlation (reference)									
ASspiMRI vs BASDAI	0.831	3.21	< 0.001	0.746	1.79	< 0.001	0.369	0.89	< 0.001
ASspiMRI vs CRP	0.675	6.93	< 0.023	0.414	4.76	< 0.205	0.636	7.86	< 0.036
ROI: SNR and CNR change									
SNR vs BASDAI	0.441	3.46	< 0.032	0.489	3.62	< 0.016	0.310	3.40	< 0.032
CNR vs BASDAI	0.248	2.07	< 0.016	0.503	1.76	< 0.008	0.431	1.92	< 0.016
Isolated examination of paravertebral lesion counts									
TC vs BASDAI	0.364	3.21	< 0.016	0.204	2.45	< 0.002	0.146	1.13	< 0.002
SP vs BASDAI	0.534	1.82	< 0.002	0.400	0.78	< 0.004	0.431	0.49	< 0.008
SIJ vs BASDAI	0.129	2.35	< 0.023	0.124	0.97	< 0.008	0.098	0.58	< 0.016
CVJ vs BASDAI	0.046	2.20	< 0.034	0.116	1.12	< 0.016	0.151	0.832	< 0.016
Modification of ASspiMRI by paravertebral lesion counts									
(ASspiMRI + TC) vs BASDAI	0.398	8.73	< 0.016	0.717	3.71	< 0.008	0.612	1.69	< 0.016
(ASspiMRI + SP) vs BASDAI	0.935	2.56	< 0.001	0.887	1.39	< 0.001	0.636	1.12	< 0.001
(ASspiMRI + TC) vs CRP	0.716	9.03	< 0.013	0.289	5.05	< 0.388	0.745	8.94	< 0.009
(ASspiMRI + SP) vs CRP	0.802	6.93	< 0.009	0.535	5.27	< 0.032	0.856	5.23	< 0.001

ROI: region of interest; TC: total paravertebral lesion count for all regions; SP: spinal processes; SIJ: small intervertebral joints; CVJ: costovertebral joints.

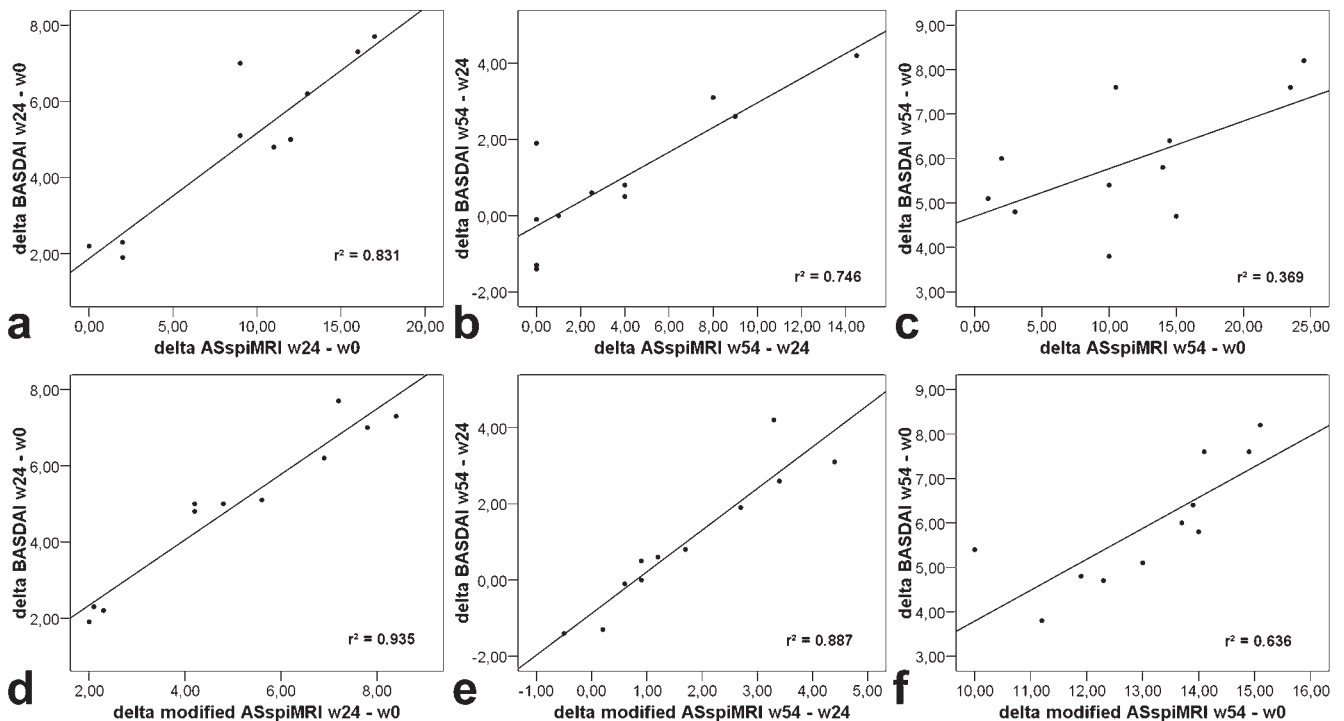


Figure 3. Linear regression analyses of correlation between changes in ASspiMRI activity score (A to C) and modified ASspiMRI (D to F) with BASDAI clinical score between Weeks 24 and 0 (A, D), Weeks 54 and 24 (B, E), and Weeks 54 and 0 (C, F). delta: difference between values for the weeks indicated.

Intraobserver variability presented significant values with 0.934 at Week 0, 0.927 at Week 24, and 0.709 at Week 54 (Table 5).

Evaluation of SNR and CNR

For 54 weeks of infliximab therapy and statistical evaluation. The SNR of a representative inflammatory lesion and the

Table 5. Interobserver and intraobserver correlation of ASspiMRI activity score, paravertebral lesion count, and measurements of contrast:noise ratio (CNR) for both readers expressed as intraclass correlation (ICC) for all 3 timepoints.

	Week 0		Week 24		Week 54	
	ICC	p	ICC	p	ICC	p
Interobserver correlation						
ASspiMRI activity score	0.856	< 0.001	0.924	< 0.001	0.751	< 0.003
Paravertebral lesion count	0.934	< 0.001	0.956	< 0.001	0.857	< 0.001
CNR	0.817	< 0.01	0.672	< 0.01	0.656	< 0.01
Intraobserver correlation						
ASspiMRI activity score	0.934	< 0.001	0.927	< 0.001	0.709	< 0.001
Paravertebral lesion count	0.941	< 0.001	0.947	< 0.001	0.908	< 0.001
CNR	0.631	< 0.001	0.451	< 0.001	0.386	< 0.01

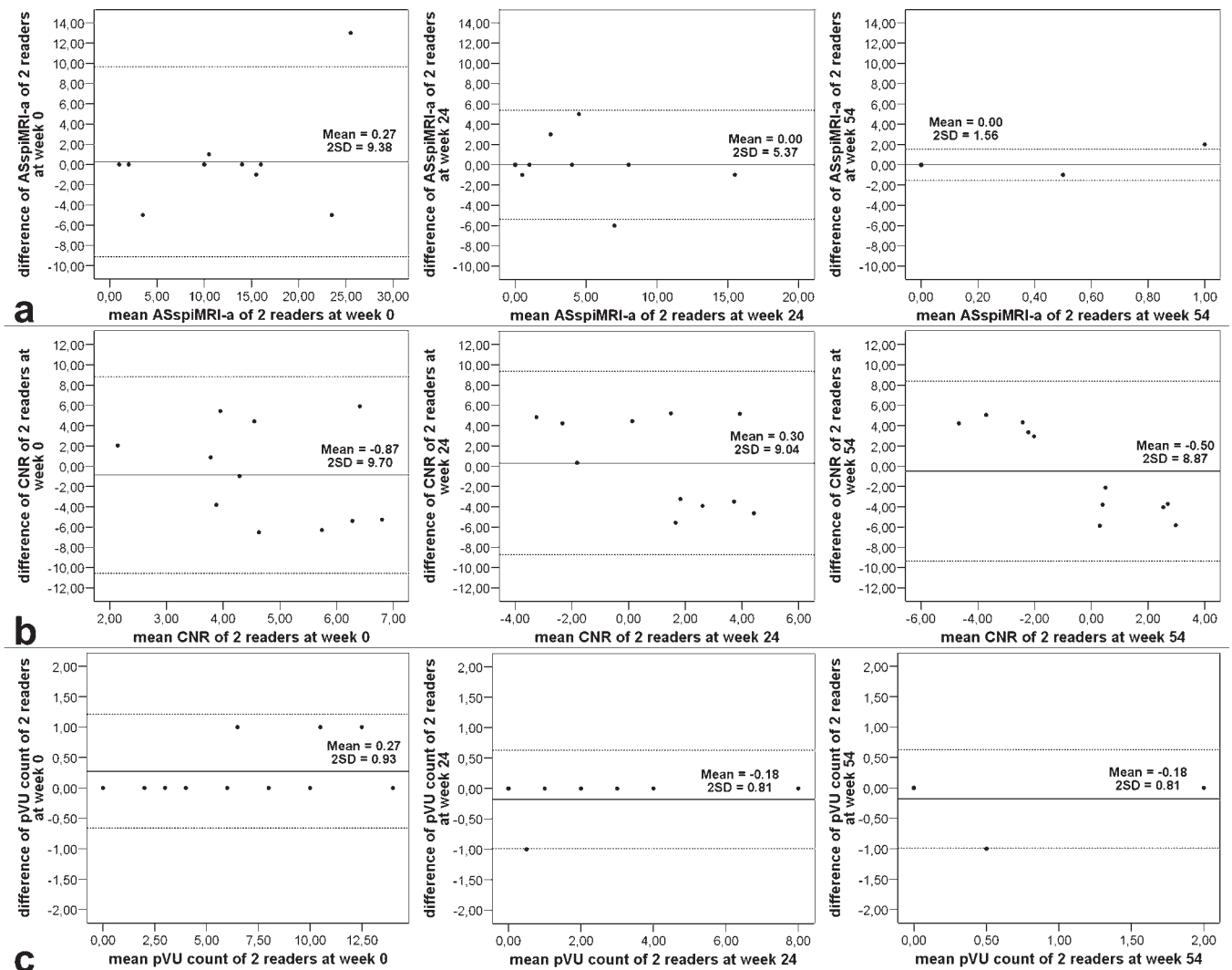


Figure 4. Bland-Altman plots illustrate interobserver variability for the results of ASspiMRI activity score (row A), measurements of contrast:noise ratio (CNR) (row B), and paravertebral lesion counts (row C) of both readers. Plots compare the results of 2 different readers and show scattering around the mean value. Center line marks the mean value, broken lines the standard deviation. Since values for ASspiMRI and the paravertebral lesion count (pVU) go to zero in Week 54, fewer than 11 dots are visible. Broad scattering around the mean value correlates with a larger interobserver variability. Scattering around the mean value is least for paravertebral lesion counts; this newly established variable therefore has considerable interobserver reproducibility. For ASspiMRI activity score, scattering was slightly increased, and was unacceptably high for measurements of CNR.

CNR of this lesion in contrast to unaffected bone marrow was determined twice by both readers for all patients. Application of SNR and CNR measurements took 10.1 minutes on average, including calculation of values. The mean SNR at study start was $9.9 (\pm 3.6)$, which decreased slightly to $5.5 (\pm 2.6)$ at Week 54. The mean CNR at study start was $4.3 (\pm 2.4)$, which decreased to $-0.8 (\pm 1.1)$ at Week 54. Both decreases were significant ($p < 0.002$ and $p < 0.001$; Table 3). Although for CNR even the improvements over time were significant ($p < 0.001$ each), they were not significant for the SNR between Weeks 54 and 24 ($p < 0.062$).

Correlation to BASDAI clinical score. The correlation coefficient of changes in SNR and BASDAI scores was 0.441 between Weeks 24 and 0 ($p < 0.032$), 0.489 between Weeks 54 and 24 ($p < 0.016$), and 0.310 between Weeks 54 and 0 ($p < 0.032$; Table 4, Figure 3). Similar values could be obtained for the CNR (0.248, $p < 0.002$; 0.503, $p < 0.008$; and 0.431, $p < 0.016$).

Interobserver and intraobserver correlation. These were calculated for the CNR only, since this value promised better correlation to clinical scores than the SNR. Results are shown in Table 5 and Figure 4.

Influence of paravertebral lesion count

For 54 weeks of infliximab therapy and statistical evaluation. The paravertebral lesion count was evaluated for each anatomical structure separately and as the total of lesions of all 3 regions (Table 3). Each reader evaluated all 3 regions twice for each patient and timepoint. Readers reported no problems for the implementation of this scoring method. Assessment of total paravertebral lesion count lasted 4.3 minutes on average. The mean paravertebral lesion count at study start was $28.4 (\pm 2.8)$, and it decreased significantly to $0.7 (\pm 0.3)$ at Week 54 ($p < 0.001$). When each region was evaluated separately, the highest lesion count at study start was detected at both the spinal processes and the costovertebral joints (2.5 and 2.9; Table 3). For all paravertebral regions a statistically significant decrease of inflammatory lesions was detected until Week 54 (Table 3). For each region there was only a slight improvement between Week 54 and Week 24 ($p < 0.25$ and $p < 0.125$).

Correlation to BASDAI clinical score. The correlation of the total and regional paravertebral lesion count is shown in Table 4. There was only a poor correlation for the total paravertebral lesion count and for the lesion count at the costovertebral and small intervertebral joints, and a good correlation for the inflammatory lesion count at the spinal processes (0.534, $p < 0.001$, between Weeks 24 and 0; 0.400, $p < 0.004$, between Weeks 54 and 24; 0.431, $p < 0.008$, between Weeks 54 and 0; Table 4, Figure 3).

Interobserver and intraobserver correlation. For ASspiMRI activity score the interobserver and intraobserver correlation

of the total and regional paravertebral lesion count showed excellent values > 0.86 and > 0.91 , respectively (Table 5). Since no significant differences between the total paravertebral lesion count and the regional lesion count could be detected, Figure 4 illustrates results only for the total count.

Evaluation of simple modifications of the ASspiMRI activity score

Since correlation of the ASspiMRI activity score to BASDAI score was good but not provable, simple modifications should be tested for a better or more reliable assessment of inflammatory activity in correlation to clinical improvements. For this purpose, both the CNR and paravertebral lesion count could be suitable. However, CNR (and SNR) measurements delivered only poor and statistically not reliable correlations to BASDAI score (data not shown) and therefore only the paravertebral lesion count was used to establish easily applicable modifications of the ASspiMRI activity score.

All paravertebral lesion counts (total and regional counts) were combined with ASspiMRI activity score in different ways (data not shown), but only simple addition of the paravertebral lesion count at the spinal processes to the ASspiMRI activity score was found to lead to a slight improvement of the correlation to BASDAI score and CRP levels, respectively (Table 4). In particular, the modified score correlated to BASDAI, with $r = 0.935$ (instead of 0.831 for the unmodified ASspiMRI activity score) between Weeks 24 and 0, $r = 0.887$ (instead of 0.746 for unmodified ASspiMRI activity score) between Weeks 54 and 24, and $r = 0.636$ (instead of 0.369 for unmodified ASspiMRI activity score) between Weeks 54 and 0. Correlation to CRP levels could be increased to $r = 0.802$ (instead of 0.675 for unmodified ASspiMRI activity score) between Weeks 24 and 0 and to $r = 0.856$ (instead of 0.636 for unmodified ASspiMRI activity score) between Weeks 54 and 0.

DISCUSSION

When considering current imaging techniques for their ability to show active inflammatory states of AS and late chronic damage at the same time, native and contrast-enhanced MRI of the spine are the only comprehensive tools that will meet both demands^{32,33,35,36} because of the excellent soft-tissue contrast and acceptable morphological depiction of bones²⁶⁻²⁸. Braun and Baraliakos and colleagues recently investigated the influence of MRI on the management of patients with early and late-stage AS^{15,29-33,37-39}. They focused not only on optimizing examination variables³³, but also on developing a promising score for exact quantification of the extent and activity of disease and chronic damage to the spine¹⁵. Because attempts to introduce a reliable radiographic score for quantification of the extent and activity of AS (such as BASRI) have failed^{25,40}, this new score may for the first time close the gap, particularly as MRI is today widely available. We think that a reliable, well tested MRI score has the potential to con-

siderably improve the quality and utility of MRI in the management of patients with AS and therefore requires multicenter investigation. The ASspiMRI activity score is the only scoring method that meets these demands and is worthy of further investigation to improve it for routine practice.

However, new scoring methods have limitations. First, their reliability has never been independently examined and so all previous reports may be biased. Second, we must allow that the ASspiMRI score specifically comprises only inflammatory and chronic lesions of the vertebral body, with no attention to other important joints of the spine such as the costovertebral and intervertebral joints and the spinal processes. Some patients may have inflammatory lesions only in these places, so this score can produce false-negative results, which might be a possible explanation for, in some cases, only small correlation with clinical changes. Third, it relies on simple imaging of features of the disease. Modern computer systems offer the opportunity to measure changes electronically, which might be more precise in the quantification of inflammatory activity. Our study was done to address these limitations and compare several contemporary algorithms and scores to find possible further improvements and to compare their performances with each other, because of the increasing demand for a reliable MRI-based score of activity of spinal inflammatory disease.

The ASspiMRI score did well in at least 3 investigative studies^{15,33,38}. The score was originally divided into 2 subgroups, the activity and chronicity scores. We focused on the activity score to quantify the activity of the disease as expressed by contrast enhancement in inflammatory lesions of the spine. In our study group only one patient had ankylosis of the spine, for which longer observational periods would be necessary to evaluate the chronicity score.

To make our data comparable with work by Braun and Baraliakos, an identical study setting was chosen. The ASspiMRI activity score did well, and was best compared with the other algorithms being studied. There was a significant reduction of the score by more than 98% after 54 weeks, and excellent correlation of the reduction with changes in the BASDAI score between Week 0 and Week 24 and between Weeks 24 and 54.

To date, the ASspiMRI score has been applied only by its originators. In our study the interobserver and intraobserver correlations were excellent, and with values of over 0.7 even better than those previously published³³, indicating that this score is easy to use and can be used successfully by even less experienced investigators, which is a basic requirement for a reliable medical test.

We examined the utility and reliability of simple counts of lesions for the costovertebral joints, the intervertebral joints, and the spinal processes. Lesions were counted and not graded further in order to simplify the procedure. This method was easy and quick. Best results and appreciable reduction rates were found at the spinal processes and the costovertebral

joints, where there were the most lesions (Figure 5). As with the ASspiMRI, the measurements of SNR and CNR and several combinations of these 3 counts and the total sum of paravertebral lesions were correlated with the BASDAI score. However, the correlations were weak for all counts including the total, with the highest values for numbers of lesions at the spinal processes. These results indicate that inflammatory lesions at this site are of special importance and show clearly that a simple count of lesions does not help to quantify the extent of the disease in this part of the spine. Some improvements have to be made to assess the degree of inflammation at this site more precisely. Nevertheless, the ASspiMRI activity score will be improved by amendment.

The score might lack sensitivity for mild inflammatory processes, so we tested the measurements of SNR and CNR to quantify a reduction or increase in inflammatory activity as an indicator of contrast enhancement. Although they are theoretically promising, these techniques were not useful. They showed a slight reduction during the 54 weeks of study, but did not correlate with the clinical scores, with correlation coefficients < 0.5. This might have been for technical reasons, as the image review software that we used had no tools to copy regions of interest for contrast:noise measurements to other images, and they had to be drawn freehand each time; the use of this measurement algorithm took about 10 minutes, which is far too long for routine practice. Even with improved software for exact transfer of the regions of interest, this algorithm may still not be correct. The interobserver and intraobserver correlations were the worst in our study, which again might be because of inadequate software. Finally, it is not possible to give an appraisal for this method based on our study data, and further studies will be needed to address these issues.

We have shown that the ASspiMRI activity score is a promising, reliable, and reproducible score for assessment of spinal inflammation in patients with AS. Although all algorithms and scores examined showed good longitudinal performance, with significant reduction, the ASspiMRI activity score correlated best with clinical changes and had the highest values for interobserver and intraobserver correlations. The reason for the difference between good longitudinal performance of all scores but varying correlation to clinical scores lies in their completely different approaches. CRP and contrast:noise ratio measurements in particular are influenced by factors other than the spinal disease. We suggest that ASspiMRI should be investigated further. The score must be amended by grading for inflammatory lesions at the paravertebral joints, particularly the spinal processes. Although promising, CNR and SNR measurements seem not to be reliable ways of quantifying inflammatory activity, as their application depends on the technical features of the PACS system used.

The limitation of our study lies in the small number of patients selected; this was necessary to achieve conditions comparable to other publications on this topic. Nevertheless,

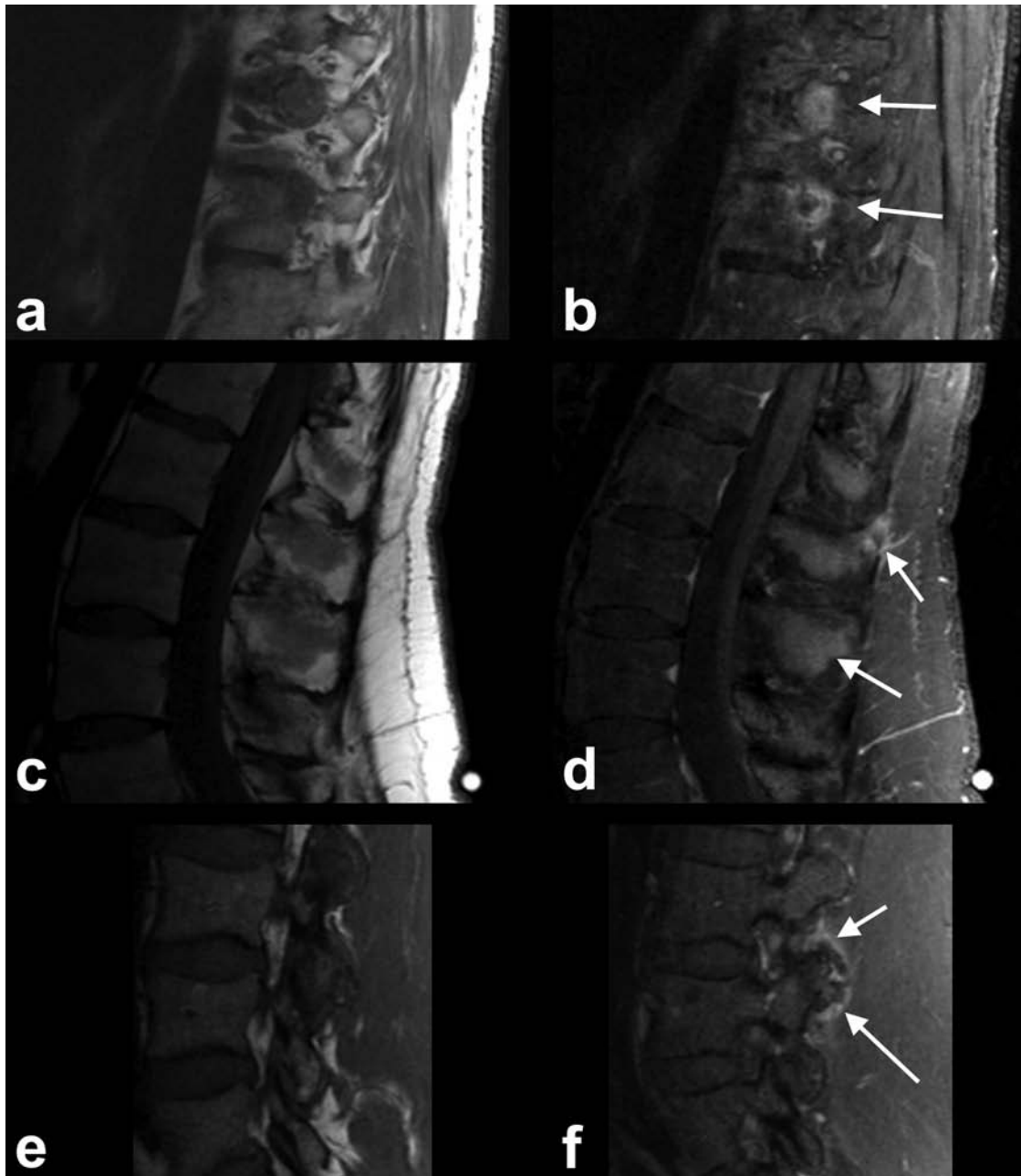


Figure 5. Paravertebral inflammatory lesions seen with MRI (sagittal planes). A, C, E: T1-weighted scans before Gd-DTPA. B, D, F: T1-weighted scans after Gd-DTPA had been given at the corresponding slice position. A and B: There is strong circular contrast enhancement of the costovertebral joints of T11 and T12 (white arrows) because the joint is involved. C and D: Laminar contrast enhancement of several spinal processes of the lumbar spine (L1–L3) caused by inflammatory involvement (white arrows). E and F: Linear contrast enhancement surrounding the small intervertebral joints (white arrows) because the joint is involved.

the values for correlation coefficients were high and significant. Further examinations with larger groups of patients are needed.

REFERENCES

1. McGonagle D, Gibbon W, O'Connor P, Green M, Pease C, Emery P. Characteristic magnetic resonance imaging enthesal changes of knee synovitis in spondylarthropathy. *Arthritis Rheum* 1998;41:694-700.
2. Braun J, Khan MA, Sieper J. Enthesitis and ankylosis in spondyloarthropathy: what is the target of the immune response? *Ann Rheum Dis* 2000;59:985-94.
3. Benjamin M, McGonagle D. The anatomical basis for disease localisation in seronegative spondyloarthropathy at entheses and related sites. *J Anat* 2001;199:503-26.
4. Resnick D, Niwayama G. Entheses and enthesopathy. Anatomical, pathological, and radiological correlation. *Radiology* 1983;146:1-9.
5. Braun J, Bollow M, Remlinger G, et al. Prevalence of spondyloarthropathies in HLA-B27 positive and negative blood

- donors. *Arthritis Rheum* 1998;41:58-67.
6. Braun J, Sieper J. The sacroiliac joint in the spondyloarthropathies. *Curr Opin Rheumatol* 1996;8:275-87.
 7. Braun J, Sieper J, Bollow M. Imaging of sacroiliitis. *Clin Rheumatol* 2000;19:51-7.
 8. Braun J, Bollow M, Sieper J. Radiologic diagnosis and pathology of the spondylarthropathies. *Rheum Dis Clin North Am* 1998;24:697-735.
 9. Wilkinson M, Bywaters EG. Clinical features and course of ankylosing spondylitis; as seen in a follow-up of 222 hospital referred cases. *Ann Rheum Dis* 1958;17:209-28.
 10. Zink A, Braun J, Listing J, Wollenhaupt J. Disability and handicap in rheumatoid arthritis and ankylosing spondylitis — results from the German rheumatological database. German Collaborative Arthritis Centers. *J Rheumatol* 2000;27:613-22.
 11. Gran JT, Skomsvoll JF. The outcome of ankylosing spondylitis: a study of 100 patients. *Br J Rheumatol* 1997;36:766-71.
 12. Ward MM. Health-related quality of life in ankylosing spondylitis: a survey of 175 patients. *Arthritis Care Res* 1999;12:247-55.
 13. Gorman JD, Sack KE, Davis JC Jr. Treatment of ankylosing spondylitis by inhibition of tumor necrosis factor α . *N Engl J Med* 2002;346:1349-56.
 14. Toussirot E, Wendling D. Current guidelines for the drug treatment of ankylosing spondylitis. *Drugs* 1998;56:225-40.
 15. Braun J, Baraliakos X, Goldner W, et al. Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab. *Arthritis Rheum* 2003;48:1126-36.
 16. Van den Bosch F, Kruithof E, Baeten D, De Keyser F, Mielants H, Veys EM. Effects of a loading dose regimen of three infusions of chimeric monoclonal antibody to tumour necrosis factor α (infliximab) in spondyloarthropathy: an open pilot study. *Ann Rheum Dis* 2000;59:428-33.
 17. Stone M, Salonen D, Lax M, Payne U, Lapp V, Inman R. Clinical and imaging correlates of response to treatment with infliximab in patients with ankylosing spondylitis. *J Rheumatol* 2001;28:1605-14.
 18. Braun J, Brandt J, Listing J, et al. Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. *Lancet* 2002;359:1187-93.
 19. van der Heijde D, Dijkmans B, Geusens 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.
 20. Baraliakos X, Listing J, Rudwaleit M, Brandt J, Sieper J, Braun J. Radiographic progression in patients with ankylosing spondylitis after 2 years of treatment with the tumor necrosis factor- α antibody infliximab. *Ann Rheum Dis* 2005;64:1462-6.
 21. Lipsky PE, van der Heijde DM, St. Clair EW, et al. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. *N Engl J Med* 2000;343:1594-602.
 22. Braun J, Golder W, Bollow M, Sieper J, van der Heijde D. Imaging and scoring in ankylosing spondylitis. *Clin Exp Rheumatol* 2002;20:S178-84.
 23. 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.
 24. 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.
 25. van der Heijde D, Spoorenberg A. Plain radiographs as an outcome measure in ankylosing spondylitis. *J Rheumatol* 1999;26:985-7.
 26. Hermann KG, Bollow M. Magnetic resonance imaging of the axial skeleton in rheumatoid disease. *Best Pract Res Clin Rheumatol* 2004;18:881-907.
 27. Hermann KG, Braun J, Fischer T, Reisschauer H, Bollow M. Magnetic resonance tomography of sacroiliitis: anatomy, histological pathology, MR-morphology, and grading. *Radiologe* 2004;44:217-28.
 28. Bollow M, Enzweiler C, Taupitz M, et al. Use of contrast enhanced magnetic resonance imaging to detect spinal inflammation in patients with spondyloarthritides. *Clin Exp Rheumatol* 2002;20 Suppl 28:167-74.
 29. Althoff CE, Hermann KG, Braun J, Sieper J. Ankylosing spondylitis — current state of imaging including scoring methods. *Z Rheumatol* 2006;65:688-99.
 30. Haraoui B. Is MRI the best method to detect response to treatment in patients with ankylosing spondylitis? *Nat Clin Pract Rheumatol* 2006;2:474-5.
 31. Maksymowych WP, Lambert RG. Spinal inflammation in ankylosing spondylitis — how and why should it be measured by MRI? *Nat Clin Pract Rheumatol* 2006;2:232-3.
 32. Braun J, Baraliakos X, Golder W, et al. Analysing chronic spinal changes in ankylosing spondylitis: a systematic comparison of conventional x-rays with magnetic resonance imaging using established and new scoring systems. *Ann Rheum Dis* 2004;63:1046-55.
 33. Baraliakos X, Hermann KG, Landewe R, et al. Assessment of acute spinal inflammation in patients with ankylosing spondylitis by magnetic resonance imaging (MRI): a comparison between contrast enhanced T1 and short-tau inversion recovery (STIR) sequences. *Ann Rheum Dis* 2005;64:1141-4.
 34. Garrett S, Jenkinson T, Kennedy LG, 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.
 35. Marzo-Ortega H, McGonagle D, O'Connor P, Emery P. Efficacy of etanercept in the treatment of the enthesal pathology in resistant spondylarthropathy: a clinical and magnetic resonance imaging study. *Arthritis Rheum* 2001;44:2112-7.
 36. Marzo-Ortega H, McGonagle D, Emery P. Etanercept treatment in resistant spondylarthropathy: imaging, duration of effect and efficacy on reintroduction. *Clin Exp Rheumatol* 2002;20 Suppl 28:175-77.
 37. 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 α receptor fusion protein etanercept. *Arthritis Rheum* 2005;52:1216-23.
 38. Baraliakos X, Landewe R, Hermann KG, 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.
 39. Baraliakos X, Braun J. Magnetic resonance imaging in spondyloarthropathies. *Joint Bone Spine* 2006;73:1-3.
 40. 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.
 41. Hanly JG, Mitchell MJ, Barnes DC, Macmillan L. Early recognition of sacroiliitis by magnetic resonance imaging and single photon emission computed tomography. *J Rheumatol* 1994;21:2088-95.