European Radiology (2018) 28:4163–4173
https://doi.org/10.1007/s00330-018-5377-3

MUSCULOSKELETAL

Whole body MRI in spondyloarthritis (SpA): Preliminary results suggest that DWI outperforms STIR for lesion detection

Frederic E. Lecouvet[1] · Nicolas Vander Maren[1] · Laurence Collette[2] · Nicolas Michoux[1] · Perrine Triqueneaux[1] · Maria Stoenoiu[3] · Frederic Houssiau[3] · Jacques Malghem[1] · Marie-Laurence Denis[1] · Ahmed Larbi[1] · Adrien Nzeusseu Toukap[3]

Received: 26 August 2017 / Revised: 1 February 2018 / Accepted: 6 February 2018 / Published online: 17 April 2018

Abstract

Purpose To compare the diagnostic accuracy of DWI and STIR sequences in Whole body (WB) MRI of SpA patients.

Materials and methods Twenty consecutive patients with confirmed active SpA and 20 controls were investigated with identical WB MRI protocols, including DWI and STIR images. Two observers recorded 'lesions' (high signal intensity foci on STIR and high b-value DWI) in 17 anatomical areas, making a 17-point 'area score' and a 40-point 'lesion score'. ROC performance, inter-observer agreement, correlation with clinical parameters and spine and sacro-iliac joints (SIJ) MRI scores were assessed.

Results SpA patients had significantly higher lesion scores on DWI than on STIR ($p<0.025$). The lesion score area under the curve was significantly higher with DWI (99.9) than with STIR (95.8, $p=0.02$). DWI lesion score ≥5 had both sensitivity and specificity ≥85 %. With STIR the best threshold ≥3 yielded sensitivity ≥85 % and specificity ≥60 %. DWI area score ≥3 yielded sensitivity ≥85 % and specificity ≥80 %. With STIR the best threshold ≥4 yielded sensitivity ≥70 % and specificity ≥80 %. Inter-observer agreement was strong for both sequences. In patients, the lesion score was positively correlated with ASDAS-CRP, log(CRP), and local MRI scores.

Conclusions DWI is a promising alternative to STIR in WB MRI to detect active SpA lesions.

Key Points

- *DWI is a robust alternative to STIR in WBMRI in SpA.*
- *DWI might be superior in discriminating relevant inflammatory and degenerative changes.*
- *Positive correlations exist between WB MRI, clinical, biological, local MRI data.*
- *Distribution and frequency of abnormal MRI findings in SpA are highlighted.*

Keywords Spondylarthritis · Spine · Whole body imaging · Diffusion, magnetic resonance imaging · Magnetic resonance imaging

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5377-3) contains supplementary material, which is available to authorized users.

✉ Frederic E. Lecouvet
frederic.lecouvet@uclouvain.be

[1] Department of Radiology and Medical Imaging, Institut de Recherche Expérimentale et Clinique (IREC),Université Catholique de Louvain (UCL), Cliniques Universitaires Saint Luc, Avenue Hippocrate, 10/2942, B-1200 Brussels, Belgium

[2] EORTC, Head of the Statistics Unit, EORTC Headquarters, Brussels, Belgium

[3] Department of Rheumatology, Cliniques Universitaires Saint-Luc, Brussels, Belgium

Abbreviations

ADC	Average diffusion coefficient
ASAS	Assessment of SpondyloArthritis International Society
ASDAS	Ankylosing Spondylitis Disease Activity Score
BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
BASFI	Bath Ankylosing Spondylitis Functional Index
CRP	C-reactive protein
DMARDs	Disease-modifying anti-rheumatic drugs
DWI	Diffusion-weighted imaging
GC	Glucocorticoids
HLA-B27	Human leucocyte antigen –B27
IBP	Inflammatory back pain
mNY	Modified New-York

NSAIDs	Non-steroidal anti-inflammatory drugs
SIJ	Sacro-iliac joints
SpA	Axial spondyloarthritis
STIR	Short tau (τ) inversion recovery
TNFa	Tumour necrosis factor alpha
T1-w	T1-weighted
VAS	Visual analogue scale
WB	Whole body

Introduction

Magnetic resonance imaging (MRI) of the sacro-iliac joints (SIJ) has been integrated as a key diagnostic criterion of axial spondyloarthropathies (SpA) according to the Assessment of SpondyloArthritis International Society (ASAS) [1–4]. Coverage of the SIJ using MRI is sufficient in most patients for a reliable diagnosis. The extension of MRI to the whole spine or to the whole body has been proposed by some researchers [5–9]. This extension may be useful in patients with non-conclusive MRI findings at the level of the SIJ; to evaluate disease activity at distance in advanced disease with ankylosed SIJ precluding evaluation of disease burden; to evaluate peripheral dominant forms of the disease; or to assess treatment response in clinical trials [10, 11].

Similar to limited SIJ and spine studies, whole body MRI studies include T1-weighted (T1-w) and STIR sequences of the whole body, sometimes complemented with spine studies [5–9]. T1-w sequences detect chronic (structural) changes in SpA, consisting of erosions, fatty transformation, bone bridging and ankylosis. STIR sequences detect active, inflammatory lesions, consisting of oedema affecting entheses, vertebral endplates, subchondral areas or the synovium [12].

Diffusion-weighted imaging (DWI) has been evaluated as a promising alternative to STIR and T2-fat saturated sequences for SIJ assessment, either quantitatively (average diffusion coefficient (ADC) measurements) or qualitatively (visual analysis of high b values DWI images) [13–17]. DWI is routinely used in whole body MRI studies performed in oncology for the evaluation of metastatic disease or haematological malignancies [18, 19]. Whether DWI could be a complement or alternative to STIR in whole body MRI studies performed in axial rheumatisms remains to be evaluated [7, 11, 20].

The purpose of our study was to compare the diagnostic accuracy of DWI and STIR sequences in whole body MRI of SpA patients.

Materials and methods

Population selection, study design and inclusion/exclusion criteria

We conducted a diagnostic accuracy study in controls and SpA patients of a DWI sequence integrated in whole body MRI studies as an alternative to the STIR sequence. The images of 20 consecutive patients (13 men, seven women; median age, 39 years) with inflammatory back pain (IBP), fulfilling the diagnostic criteria for SpA (either the modified New-York (mNY) criteria, or the Assessment of SpondyloArthritis International Society (ASAS) criteria) investigated in our department between January and December 2015 [4, 21], were evaluated for this study. One additional claustrophobic patient was referred during the same period but refused MRI and could not be included. Seven patients had newly diagnosed SpA (<1 year), six had had SpA for 1–4 years and seven had had SpA for >4 years (Table 1).

Disease activity was clinically assessed by the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Visual Analogue Scale (VAS), Bath Ankylosing Spondylitis Functional Index (BASFI) and Ankylosing Spondylitis Disease Activity Score (ASDAS). On a 10-point scale, a BASDAI score ≥4 indicates active disease. ASDAS values <1.3 indicate no activity (remission); values of 1.3–2.1, 2.1–3.5 and >3.5 indicate moderate, high and very high activity, respectively [22]. The patient's C-reactive protein (CRP, normal value <5 mg/L) was measured to assess the biological activity and for ASDAS-CRP scoring. HLA-B27 was available in all but one patient. Whole body MRI was requested for first evaluation or for disease re-evaluation triggered by clinical or biological signs of disease evolution.

Whole body MRI studies included T1-w, STIR and DWI sequences in addition to our department's routine SpA MRI protocol that covers the whole spine with sagittal T1-w and STIR sequences and the SIJ with oblique coronal T1-w and T2 fat-saturated sequences. Limited spine and SIJ 'Berlin' MRI scores (12-point SIJ score and 69-point spine score) were calculated as done routinely in our unit [23, 24].

Twenty controls (19 men and one woman, median age, 59 years, range: 25–73) were selected among patients referred by the oncology department for cancer staging who had been investigated with identical whole body MRI studies. Controls had to be younger than 75 years with a negative whole body MRI for bone and other organs disease (to avoid identification of controls during readings), and no inflammatory rheumatological disorders (based on a questionnaire the day of examination and careful review of their medical records and biological parameters). The study was approved by our institution's ethics committee, which did not require informed consent for the retrospective review of prospectively acquired data.

Whole body MRI protocol, target areas and reading

Whole body MRI studies were acquired on a 3T magnet (Philips Ingenia, Best, The Netherlands), covering the body from the vertex to mid-thighs and from the anterior to the

Table 1 Demographics/patients characteristics

Parameters	20 SpA patients	20 controls
Gender	13 men, 7 women	19 men, 1 woman
Age (y)	39 (19–57)	59 (25–73)
Disease duration (y)	8.4 (1–36)	-
Symptoms: IBP (n)	20	-
Radiographic/non-radiographic disease	17/3	
VAS (score 0–10)	5.8 (0–10)	-
BASDAI (score 0–10)	5.1 (1.9–9.4)	-
BASFI (score 0–10)	4.1 (0.2–7.1)	-
ASDAS-CRP (score)	3.1(0.9–4.8)	-
CRP (mean, range mg/ml)	11.3(0.5–50)	All < 0.5
HLA-B27 (+/-) (available in 19)	15 / 4	-
NSAIDs (yes/no)	20 / 0	-
DMARDs alone (yes/no)	1 / 19	-
DMARDs + GC (yes/no)	2 / 18	-
TNFα blockers (yes/no)	3 / 17	-

pos/neg positive/negative, *ASAS* Assessment of SpondyloArthritis international Society, *ASDAS* Ankylosing Spondylitis Disease Activity Score, *BASDAI* Bath Ankylosing Spondylitis Disease Activity Index, *BASFI* Bath Ankylosing Spondylitis Functional Index, *CRP* C-reactive protein, *DMARDs* disease modifying anti-rheumatic drugs,

GC glucocorticoids, *HLA-B27* Human Leucocyte Antigen-B27, *IBP* inflammatory back pain, *mNY* modified New-York,

NSAIDs non-steroidal anti-inflammatory drugs, *TNFa* tumour necrosis factor alpha, *VAS* visual analogue scale

posterior thoraco-abdominal wall. Sequence parameters are detailed in Table 2.

The presence or absence of involvement according to high signal intensity zones ('lesions') on STIR and DWI sequences was recorded in 17 anatomical areas (see Table 3). A 'lesion' was defined as a high signal intensity area consistent with bone marrow oedema involving vertebral angles, endplates, costo-vertebral or facet joints, anterior or posterior vertebral ligaments; subchondral areas of acromio-clavicular, sterno-clavicular, manubrio-sternal and chondro-sternal joints, pubic symphysis; or peripheral entheses, i.e. ischiatic tuberosities, greater and lesser tuberosities at the level of hips and shoulders, and acromions.

The following observations were not considered positive for lesions: vertebral foci centrally located within vertebral bodies, supposed to correspond to vertebral haemangiomas; vertebral foci underlying vertebral endplate deformities, supposed to

Table 2 Whole body MRI protocols: detailed parameters on a 3T magnet (Philips Ingenia 3 T)

Parameters	T1-weighted	STIR	DWI
Plane	Coronal	Coronal	Axial
No. of sections/thickness (mm)/gap (mm)	60/4/0.4	60/4/0.4	25/5/0.5
Field of view	450 × 279	450 × 279	430 × 366
Acquisition matrix	452 × 229	380 × 182	83 × 122
Number of stations	4	4	4
Phase encoding	FH	FH	AP
IPAT factor	2	0	2
Phase oversampling	2 × 140 mm	2 × 140 mm	0
Repetition time/echo time/turbo factor (ms)	643/7.6/3	4,432/13/30	4,500/68/5
Flip angle (°)	90	90	130
Fat suppression technique	-	IR (IT = 200 ms)	STIR
b values (s/mm^2)	-	-	50, 1000
Approximate duration (min:s)	5:20	12:20	13:20

STIR short tau inversion recovery, *DWI* diffusion-weighted imaging, *IPAT* integrated parallel acquisition techniques, *IR* inversion recovery, *IT* inversion time, *FH* feet-head, *AP* anteroposterior

Table 3 Reading sheet for record of involved anatomical areas, and scoring

Anatomical areas	Lesion	N foci (spine) or slices (SIJ)	Lesion score	Rating**
	0/1*			
Cervical spine	0 - 1	x	0 - 3	N foci : 0=0/1=1/2=2 /3= ≥ 3
Thoracic spine	0 - 1	x	0 - 5	N foci: 0=0/1=1/2=2-5/3=6-10/4=11-15/5=≥15
Lumbar spine	0 - 1	x	0 - 5	N foci: 0=0/1=1/2=2-5 /3=6-10/4=11-15/5=≥15
Sterno-clavicular joints	0 - 1		0 - 2	1 if unilateral involvement, 2 if bilateral
Manubrio-sternal joint	0 - 1		0 - 1	
Chondro-sternal joints	0 - 1		0 - 2	1 if unilateral involvement, 2 if bilateral
Right sacro-iliac joint (SIJ)	0 - 1	x	0 - 5	N slice(s) : 0=0/1=1/2=2-5/3=6-10/4=11-15/5=≥15
Left SIJ	0 - 1	x	0 - 5	N slice(s) : 0=0/1=1/2=2-5/3=6-10/4=11-15/5=≥15
Pelvis (Ischio; symphysis)	0 - 1		0 - 4	1 per ischion ; 2 for symphisis (1 per side)
Right hip entheses	0 - 1		0 - 1	
Left hip entheses	0 - 1		0 - 1	
Right hip joint	0 - 1		0 - 1	
Left hip joint	0 - 1		0 - 1	
Right shoulder entheses	0 - 1		0 - 1	
Left shoulder entheses	0 - 1		0 - 1	
Right shoulder joint	0 - 1		0 - 1	
Left shoulder joint	0 - 1		0 - 1	
	Area score 0–17		**Lesion score 0–40**	

*Lesion is defined as an abnormal focus of high SI on STIR and DWI

**Quantitative information was recorded in the most extensive and frequently involved areas, spine and SIJ. The number of slices (SIJ) or abnormal foci (spine) were recorded, and converted into grading on a 3-point (cervical) or 5-point (thoracic and lumbar) scale. Other anatomical areas were given 1 point if involved, except for the pubis, which was given 2 points (1 per side).

Example of use for lumbar spine: a single focus was given a score of '1'; a number of foci between 6 and 10 within the lumbar spine was given a score '3' for this segment

Example of use for the SIJ: a single focus was given a score '1'; a number of slices showing abnormal signal between 6 and 10 was given a score '3' for this SIJ

correspond to recent vertebral compression fractures or Schmorl nodes. Joint and entheses located more peripherally than hips and shoulders were not evaluated, as these peripheral areas are not covered by routine whole body MRI examinations.

Two observers (one senior radiologist and one radiology fellow with 10 and 3 years of experience in whole body MRI reading, respectively) read the images separately, blinded to whether the images were from patients or controls. DWI images were reconstructed as multi-planar reformats in the coronal plane exactly matching the native STIR images. STIR and DWI sequences were read separately in a random order at >4 weeks' interval on PACS workstations. Each observer assessed the involvement of each individual anatomic area, resulting in an 'area score' ranging from 0 to 17 representing the number of involved regions identified by high signal intensity foci. The number of abnormal foci in each spine segment was recorded as well as the number of slices showing abnormal signal intensity in each SIJ.

Based on the numbers of foci in the spine and slices in the SIJ, a score was provided to each anatomical area (Table 3) and the sum of the scores was calculated. This 'lesion score' was obtained to reflect the number of abnormally high signal intensity zones (lesions), beyond the number of involved anatomical areas. This second score first aimed at evaluating and differentiating the severity of abnormalities between patients and controls, and at defining quantitative distinctive thresholds. This quantitative index could be correlated with biological and clinical parameters of disease activity and with spine and SIJ 'Berlin' MRI scores [23, 24]. The lesion score (range 0–40) was defined by the sum of the individual area score within each area counting 1 for each single area involved and the number of abnormal foci/slices for each spine segment/SIJ (Table 3). Inter-observer agreement based on area and lesion scores and correlation of the area and lesion scores with clinical and biological parameters and with local spine and SIJ MRI scores were assessed.

Statistical analysis

Continuous variables were described using boxplots and compared between groups by the Wilcoxon test and within patients by the signed rank test [25, 26]. Inter-observer

diagnostic agreement was assessed by the overall percent agreement. Discrimination was measured by receiver operating characteristic (ROC) curve analysis, using clinical diagnosis as a reference standard. To avoid bias, the best performer's areas under curve (AUCs) for each sequence were compared using McNemar's test. Linear regression was used in patients to measure the association with biological or known imaging parameters. To adjust for multiplicity, a significance level of <0.025 was used whenever the same test was performed for both observers, and a significance level of <0.01 was used when testing associations with the five clinical parameters. All tests and confidence intervals were 2-sided. All analyses were post hoc. No formal power calculations were performed. Statistical analysis software version 9.2 (SAS®, Bethesda, MD, USA) was used. The optimal threshold score to identify patients was determined for each sequence by inspection of waterfall plots and ROC curves.

Results

Anatomical distribution and frequencies of DWI and STIR abnormalities

The anatomical area involvement is illustrated in Fig. 1 (and Supplementary Online Table 1). The profiles of anatomical involvement (Supplementary Online Fig. 1) suggest distinctive features of involved areas: SIJ, cervico-thoraco-lumbar spine and hip entheses were most commonly involved in patients, spine and hip entheses in controls. Some areas were strongly indicative of SpA, as they were not (manubrio-sternal, chondro-costal) or were only marginally (hip synovitis, ischiatic tuberosities and pubis grouped in 'pelvis') involved in controls (Supplementary Online Fig. 1). This was consistent across sequences and observers.

Comparison of number of involved areas and lesion scores between patients and controls

Both the 17-point area score and the 40-point lesion score were significantly higher in patients than in controls, for both observers and with both sequences (see Table 4, $p<0.0001$ to $p=0.0019$).

No difference in clinical or biological data, number of involved areas and lesion scores was observed between patients with radiographic and non-radiographic SpA.

The number of involved areas was substantial in controls, which likely corresponds to a relatively high prevalence of degenerative changes. Indeed, only five controls had 'fully negative' MR studies using STIR, and only nine using DWI sequence (Supplementary Online Fig. 1). However, controls presented much lower numbers of active foci than patients within each individual anatomical area, especially in the spine and SIJ, and thus had lower lesion scores.

In controls, both observers identified a median of one more anatomical area with STIR than with DWI (median 2 vs. 1, $p=0.0032$ and median 2 vs. 1, $p=0.0106$ for Observers 1 and 2, respectively). In patients, the number of areas seen on DWI versus STIR did not differ significantly (6 vs. 5, $p=0.0576$ and 4.5 vs. 4.5, $p=0.9986$, Table 4). In patients, the lesion score was significantly higher with DWI than with STIR for both reviewers (median 12.5 vs. 7, median difference +5 with DWI, $p<0.0001$ for Observer 1; median 9.5 vs. 8, median difference +2, $p=0.0053$ for Observer 2). Figures 2 and 3 illustrate one typical patient and one control.

Diagnostic performance and thresholds to identify patients based on area and lesion scores (Figs. 4 and 5)

The area score had lower diagnostic performance than the lesion score (Fig. 5), in particular with STIR.

With STIR the area under the ROC curve (AUC) for the area score was 84.0 % (95 % CI: 72.1–95.9) for Observer 1 and 80.7 % (95 % CI: 67.2–94.1) for Observer 2. With DWI the AUC for the area score was 96.0 % (95 % CI: 91.1–100) for Observer 1 and 92.9 % (95 % CI: 85.7–100) for Observer 2.

For the lesion score, using STIR the AUC was 95.9 % (95 % CI: 90.9–100) for Observer 1 and 86.4 % (95 %CI: 75.3–95.7) for Observer 2. With DWI the AUC of the lesion score was 99.9 % (95 % CI: 99.5–100) for Observer 1 and 92.0 % (95 % CI: 82.6–100) for Observer 2.

Comparing the best achieved AUC of each sequence (i.e. that of Observer 1) on the area score, the AUC was higher for DWI than STIR (96 vs. 84, +12.0, CI: +1.6 to + 22.4, $p=0.0233$). Comparing best observers for the lesion score (i.e. Observer 1), DWI had only a marginally higher AUC than STIR (99.9 vs. 95.9, +4.0, CI: -0.8 to +8.8, $p=0.1028$, Fig. 5b) as both AUCs were close to 1 (Fig. 5).

Regarding the lesion score, the optimal STIR threshold (≥3/40) yielded 85 % (85 %) sensitivity and 90 % (60 %) specificity for Observer 1 (Observer 2).

Regarding the lesion score, the optimal DWI threshold (≥5/40) yielded 95 % (85 %) sensitivity and 100 % (90 %) specificity for Observer 1 (Observer 2).

Regarding the area score, the optimal STIR threshold (≥4/17) yielded 70 % (75 %) sensitivity and 80 % (80 %) specificity for Observer 1 (Observer 2).

Regarding the area score, the optimal DWI threshold (≥3/17) yielded 100 % (85 %) sensitivity and 80 % (80 %) specificity for Observer 1 (Observer 2).

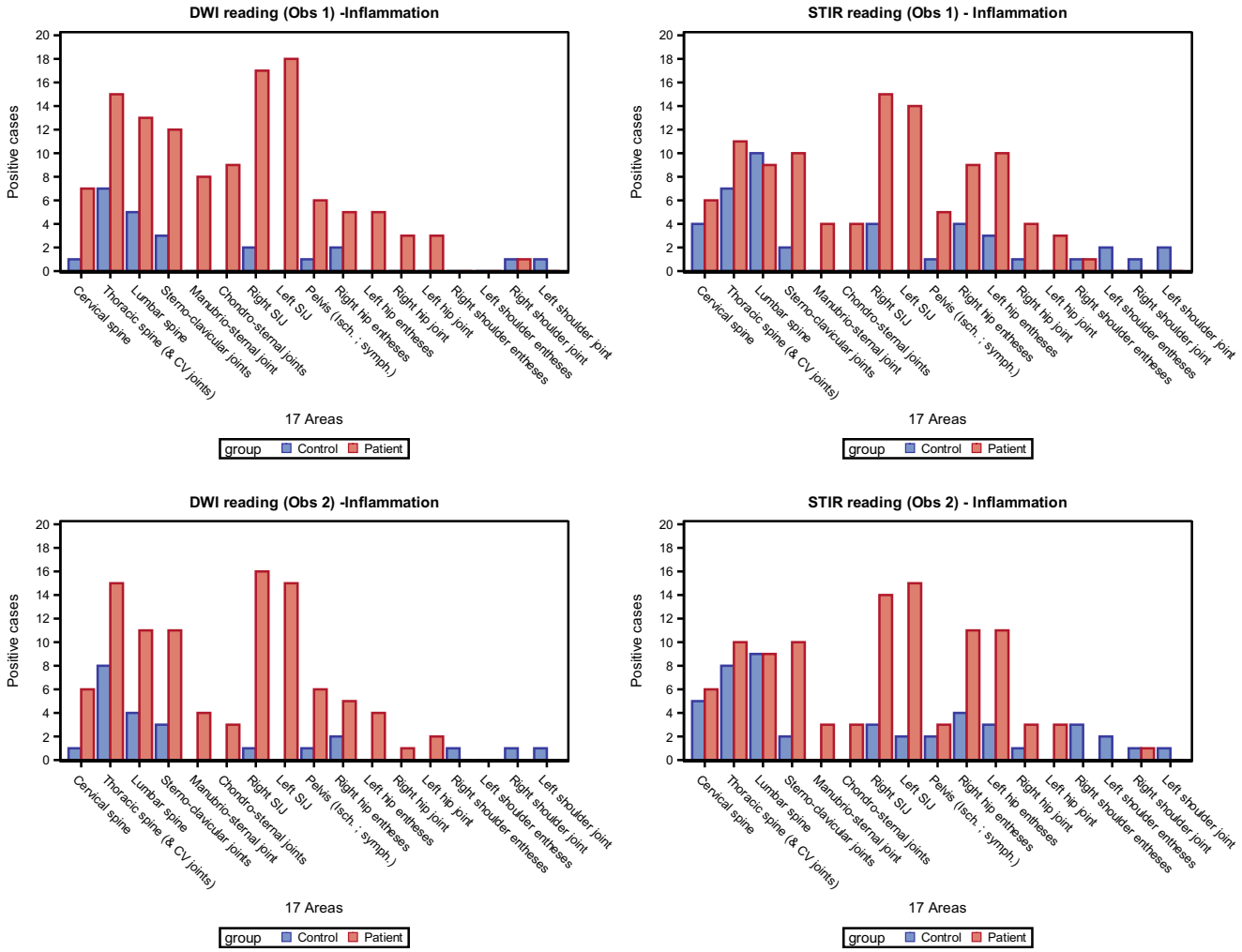

Fig. 1 Distribution of involved areas by sequence and observer. Schematic view of the distribution of positive areas for both observers on STIR and DWI sequences, in patients (red bars) and controls (blue bars). The graph shows the number of patients and controls for whom each area is positive. Sacro-iliac joints and spine represent most frequently involved sites in spondyloarthritis (SpA). Manubrio-sternal and chondro-sternal joints are highly specific for SpA. The spine is frequently involved with areas of abnormal signal, both in patients and controls (thought to represent inflammation in patients, juxta discal degenerative changes in controls)

Inter-observer agreement

When assessing the presence of abnormalities in the 17 areas of the 20 patients, the two observers very strongly agreed using DWI (Agreement: 92.7 %, CI: 89.9–95.4) and STIR (95.6 %, CI: 93.4–97.8). For controls, the agreement was 95.9 % using DWI (CI: 93.8–98.0) and 95.9 % using STIR (CI: 93.4–97.8).

Defining 'patients' by a lesion score ≥5 on DWI, the observers agreed on 90 % of their diagnostics (36/40, CI: 76.3–96.2). Defining them by a lesion score ≥3 with STIR to identify patients, they agreed on 85 % of the diagnostics (34/40, CI: 70.2–94.3).

Defining 'patients' by an area score ≥3 on DWI, the observers agreed on 87.5 % of their diagnostics (35/40, CI: 73.2–95.8). Defining them using an area score ≥4 on STIR, both observers agreed in all but one case (97.5 %, 39/40, CI: 86.8–99.9).

Correlations with clinical and biological parameters

We observed a positive correlation for both observers and both sequences between the lesion score and ASDAS-CRP (borderline with DWI, Observer 1, p=0.0166; other p values <0.01) and between DWI readings and logCRP values (p values <0.01, except for STIR Observer 1 p=0.025; Supplementary Online Fig. 2 and Supplementary Online Table 2). The area score did not correlate with clinical and biological parameters. A positive correlation was also observed between all lesion and area scores and the limited SIJ and spine 'Berlin' MRI scores (all p values <0.01) (Supplementary Online Table 2).

Table 4 Areas and lesion scores on DWI and STIR readings by Observers 1 and 2

	Area score / 17			Lesion score / 40		
	Controls (N=20)	Patients (N=20)	*p*-value[1] (Wilcoxon)	Controls (N=20)	Patients (N=20)	*p*-value[1] (Wilcoxon)
DWI (Obs 1)						
Median	1.0	6.0	<0.0001	0.0	12.5	<0.0001
Range	0.0–4.0	3.0–13.0		0.0–4.0	4.0–31.0	
Mean (SD)	1.15 (1.39)	6.10 (2.85)		0.45 (1.10)	13.35 (7.45)	
STIR (Obs 1)						
Median	2.0	5.0	0.0007	0.0	7.0	<0.0001
Range	0.0–6.0	1.0–12.0		0.0–5.0	2.0–24.0	
Mean (SD)	2.10 (1.83)	5.25 (2.69)		0.90 (1.45)	8.35 (6.05)	
Within-patient difference DWI minus STIR (Obs 1)						
Median	-1.0	1.0		0.0	5.0	
Range	-4.0–2.0	-3.0–4.0		-4.0–3.0	-1.0–13.0	
Mean (SD)	-0.95 (1.64)	0.85 (1.76)		-0.45 (1.39)	5.00 (3.51)	
p-value[2]	0.0032	0.0576		0.1445	<0.0001	
DWI (Obs 2)						
Median	1.0	4.5	<0.0001	1.0	9.5	<0.0001
Range	0.0–4.0	1.0–9.0		0.0–9.0	0.0–23.0	
Mean (SD)	1.15 (1.27)	4.95 (2.31)		1.90 (2.27)	11.60 (6.85)	
STIR (Obs 2)						
Median	2.0	4.5	0.0019	2.0	8.0	<0.0001
Range	0.0–6.0	1.0–12.0		0.0–7.0	1.0–24.0	
Mean (SD)	2.30 (1.75)	5.10 (2.79)		2.25 (1.97)	8.65 (6.10)	
Within-patient difference DWI minus STIR (Obs 2)						
Median	-1.0	0.0		0.0	2.0	
Range	-4.0–2.0	-6.0–3.0		-4.0–4.0	-3.0–10.0	
Mean (SD)	-1.15 (1.69)	-0.15 (2.35)		-0.35 (2.23)	2.55 (3.49)	
p-value[2]	0.0106	0.9986		0.5471	0.0053	

1 Wilcoxon test for testing a difference between controls and patients

2 Wilcoxon Signed rank test for testing a mean difference of 0 between DWI and STIR within patients. A difference >0 indicates a larger number found with DWI compared to STIR. The significance level is 0.025 in this table

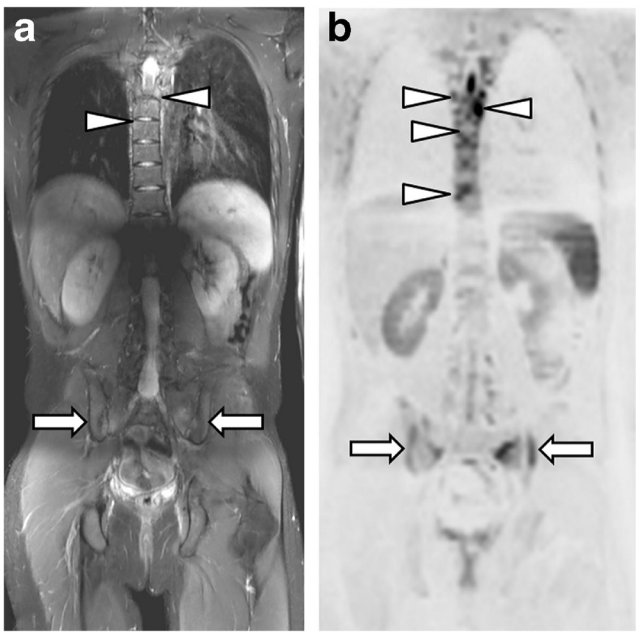

Fig. 2 43-year old male SpA patient presenting with inflammatory pelvic pain. Comparison of STIR (**A**) and DWI (inverted grey scale) (**B**) whole body magnetic resonance imaging (MRI) sequences. Involvement of the thoracic spine (central and lateral vertebral endplates, costo-vertebral junctions) (arrowheads) and SIJ (arrows) is visible on both images but much more evident on DWI, due to background signal suppression

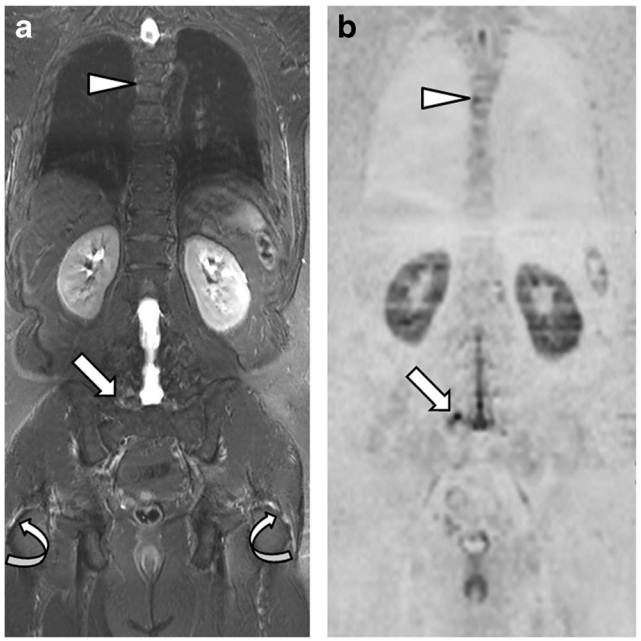

Fig. 3 63-year-old male with negative WB MRI screening for prostate cancer metastases, included in the control population. Comparison of STIR (**A**) and DWI (inverted grey scale) (**B**) whole body MRI sequences. Degenerative involvement of the thoracic spine (arrowheads) and of a right posterior (zygapophyseal) joint within the lumbar spine involvement (arrows) is visible on both images. Note that involvement of tendinous insertions on the greater trochanter is visible only on STIR images (curved arrows in A), not on DWI

Discussion

The primary objective of this study was to compare the use of DWI and STIR sequences as part of whole body MRI studies to detect active foci in SpA patients. Our results show that DWI is a robust alternative to the STIR sequence in whole body MRI studies obtained for active inflammatory lesion detection in SpA patients. Our study also provides insights into the distribution, frequency and severity of abnormal findings in SpA patients in comparison to a control population and suggests distinctive patterns between these populations. Finally, we highlight some correlations between whole body MRI findings and clinical and biological parameters in active SpA, and with local spine and SIJ MRI scores.

Compared to STIR, DWI detected a smaller number of involved areas in controls and provided higher lesion scores in SpA patients. Comparing the area score and lesion scores, DWI was more discriminant than STIR in distinguishing patients from controls based on AUC. This suggests that DWI could be useful in whole body MRI protocols as it offers higher sensitivity for the detection of inflammatory lesions compared to STIR images. Training may be necessary, as the superiority of DWI over the STIR sequence was more evident for the experienced reader. The superiority of DWI over STIR may relate to its higher signal to background ratio, already reported in oncological whole body protocols using DWI [27]. The ability to differentiate inflammation from degenerative changes represents a potential added value for DWI. This value of DWI to detect active inflammation in SpA patients was highlighted in preliminary works targeting the SIJ [13, 14]. The underlying mechanism of signal alteration may be that inflammatory lesions with infiltration by specific cells and alterations within the extracellular compartment are expected to induce changes in the intra-/extracellular balance and thus signal alterations [20]. Bone marrow oedema and infiltration of the normal bone marrow by inflammatory cells are thought to induce a local increase in water movement, resulting in increased local diffusion and higher ADC values within the lesions. Bozgeyik et al. used quantitative DWI to show that patients with inflammatory sacroiliitis had significantly higher ADC values measured from iliac and sacral bones compared to patients with back pain of mechanical origin. DWI had a similar sensitivity compared to that of gadolinium enhanced T1-weighted images [13]. Gezmis demonstrated increased ADC values in early stages of SpA, and suggested a return to normal as an interesting marker of response [14]. Sanal et al. using qualitative DWI sequences showed that active inflammation could be detected on high b value DWI images of the SIJ, challenging STIR and contrast-enhanced fat saturated T1-w sequences [15, 16].

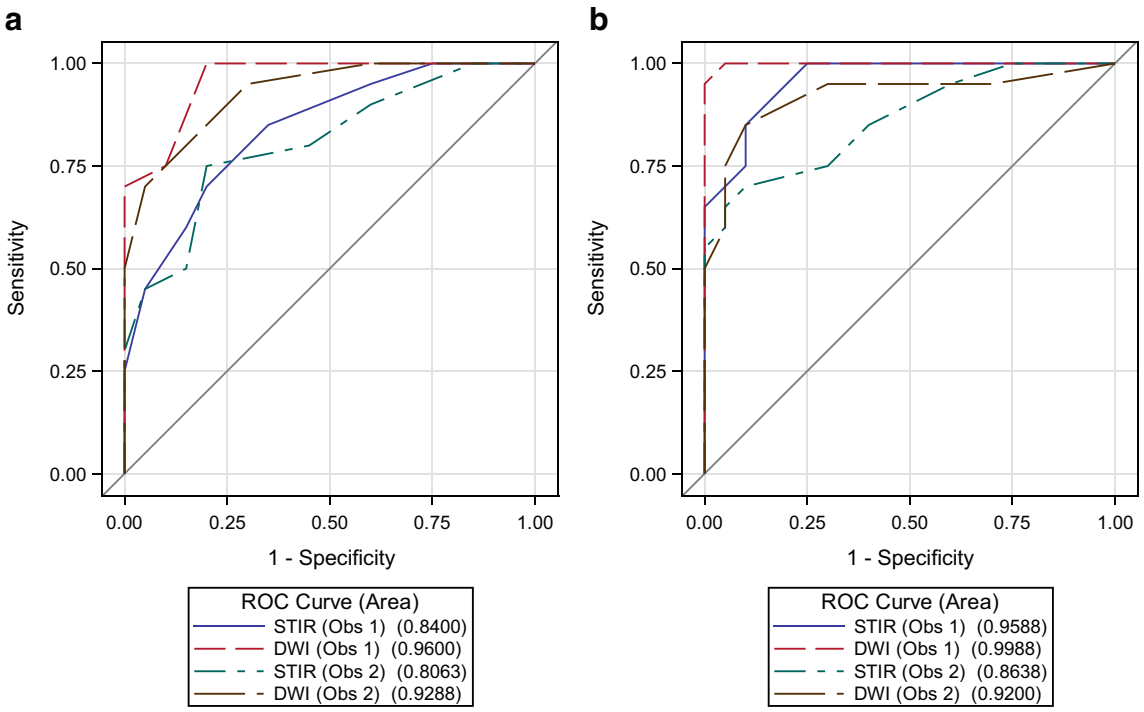

Fig. 4 Receiver operating characteristic (ROC) curve showing the performance of the diagnostic tests for both observers (obs) (coloured lines); the diagonal line shows no discrimination, the best performance is achieved with sensitivity and specificity of 100 (top left corner). (**a**) ROC curves for diagnosing patients using the area score by DWI and STIR for obs. 1 and 2. *STIR:* Area under the ROC (AUC)=84.0, 95 % CI: 72.1–95.9 for obs. 1 and AUC=80.7, 95 % CI: 67.2–94.1 for obs. 2. *DWI:* AUC= 96.0, 95 % CI: 91.1–100 for obs. 1 and AUC= 92.9, 95 % CI: 85.7–100 for obs. 2. (b) ROC curves for diagnosing patients with the lesion score by DWI and STIR for obs. 1 and 2. *STIR*: AUC=95.9 %, 95 % CI 90.9–100 for obs. and AUC=86.4 %, 95 % CI: 75.3–95.7 for obs. 2. DWI: AUC= 99.9 %, 95 % CI: 99.5–100 for obs. 1 and AUC= 92.0 %, 95 % CI: 82.6–100 for obs. 2

Second, the current study highlighted that some anatomical areas were strongly suggestive of involvement by SpA. This was the case for manubrio- and chondro-sternal joint involvement and for hip synovitis, which were rarely observed in controls. In contrast, the spine and SIJ were involved in both patients and controls. The most commonly areas involved in controls were the spine, likely representing degenerative 'Modic 1' changes; the SIJ, also representing degenerative changes; and enthesitis of the proximal femoral trochanters, another very common observation of degenerative origin. Therefore, MR examinations should be interpreted cautiously when only limited changes are detected in these areas. This parallels the caution suggested in the current guidelines that require at least two separate foci or two slices showing the same focus at the level of the SIJ, or more than three vertebral foci, to consider that an MRI examination is relevant for inflammation and diagnostic for SpA [6, 28, 29].

Third, we observed a positive correlation between both DWI and STIR lesion scores derived from whole body MRI and logCRP levels and ASDAS-CRP. No correlation was found with other clinical scores, nor between the number of involved anatomical areas at MRI and any biological parameter. Of note, a positive correlation between quantitative DWI data and CRP levels was reported in a previous study [30]. Limited 'Berlin' MRI scores derived from spine and SIJ readings also correlated with the lesion and area scores for both observers and the Berlin SIJ score correlated with ASDAS-CRP. These positive correlations support the value of the area and lesion scores proposed in our study.

This study has several limitations. First, we did not assess the value of T1-w MR images for detecting SpA lesions or the value of ADC measurements for lesion detection, as this has been done in studies comparing inflammatory and degenerative changes in limited areas [13, 14]. Second, the ages of our control and patient populations were different because selection of controls was based on negative whole body MRI studies and no rheumatological disease. An aged-matched control group may have presented less degenerative changes than ours, which could have resulted in stronger discriminating features between patients and controls. Third, our whole body MRI protocols did not cover the body from head to toes as the knee-to-toes area is not covered in cancer patients. The respective value of DWI and STIR sequences in the diagnosis of peripheral entheses and

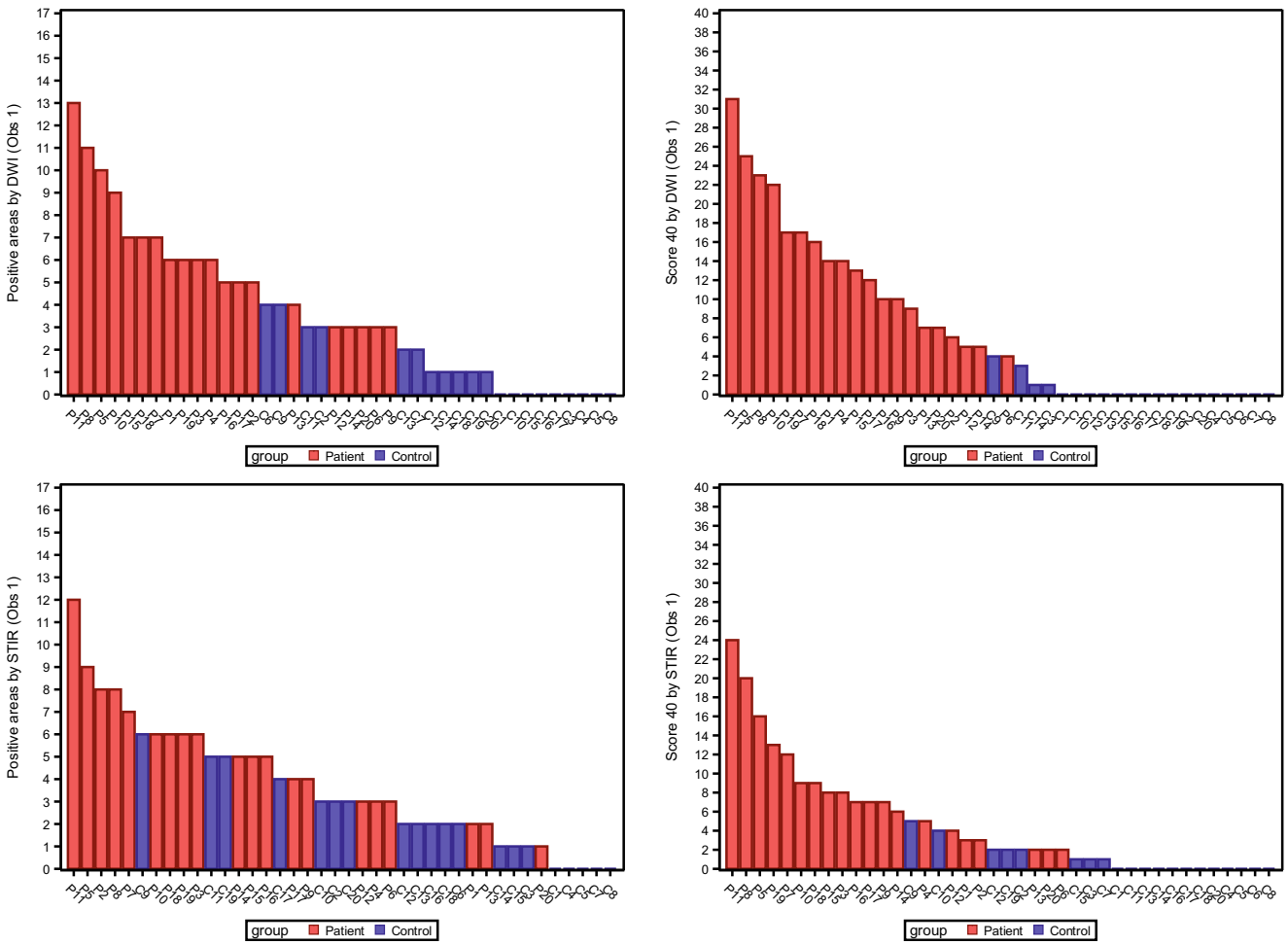

Fig. 5 Waterfall plots illustrating the area score s (/17) and lesion score (/40) in patients (red bars) and controls (blue bars) for Observer 1

joints involvement remains to be assessed [31]. Our study does not address the diagnostic accuracy of DWI and STIR for SpA compared to other arthropathies, but only for AS compared to controls; further studies of diagnostic accuracy using AS patients and controls with other clinically confirmed arthropathies are necessary to determine this.

In conclusion, our preliminary study suggests that DWI is promising as a part of whole body MRI studies performed for the detection of active foci in SpA and might be superior to STIR for the detection of lesions and discrimination between relevant inflammatory changes and degenerative changes observed in a control population.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is F.E. Lecouvet.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent The institution's ethics committee did not require informed consent for the retrospective review of prospectively acquired data.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- retrospective
- case-control study
- performed at one institution

References

1. Braun J, Sieper J (2007) Ankylosing spondylitis. Lancet 369: 1379–1390
2. Rudwaleit M, van der Heijde D, Khan MA, Braun J, Sieper J (2004) How to diagnose axial spondyloarthritis early. Ann Rheum Dis 63:535–543
3. Mandl P, Navarro-Compan V, Terslev L et al (2015) EULAR recommendations for the use of imaging in the diagnosis and management of spondyloarthritis in clinical practice. Ann Rheum Dis 74:1327–1339
4. Rudwaleit M, van der Heijde D, Landewe R et al (2009) The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 68:777–783
5. Rennie WJ, Dhillon SS, Conner-Spady B, Maksymowych WP, Lambert RG (2009) Magnetic resonance imaging assessment of spinal inflammation in ankylosing spondylitis: standard clinical protocols may omit inflammatory lesions in thoracic vertebrae. Arthritis Rheum 61:1187–1193
6. Hermann KG, Baraliakos X, van der Heijde DM et al (2012) Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. Ann Rheum Dis 71:1278–1288
7. Althoff CE, Appel H, Rudwaleit M et al (2007) Whole-body MRI as a new screening tool for detecting axial and peripheral manifestations of spondyloarthritis. Ann Rheum Dis 66:983–985
8. Mager AK, Althoff CE, Sieper J, Hamm B, Hermann KG (2009) Role of whole-body magnetic resonance imaging in diagnosing early spondyloarthritis. Eur J Radiol 71:182–188
9. Weber U, Maksymowych WP, Jurik AG et al (2009) Validation of whole-body against conventional magnetic resonance imaging for scoring acute inflammatory lesions in the sacroiliac joints of patients with spondylarthritis. Arthritis Rheum 61:893–899
10. Karpitschka M, Godau-Kellner P, Kellner H et al (2013) Assessment of therapeutic response in ankylosing spondylitis patients undergoing anti-tumour necrosis factor therapy by whole-body magnetic resonance imaging. Eur Radiol 23:1773–1784
11. Lecouvet FE, Michoux N, Nzeusseu Toukap A et al (2015) The Increasing Spectrum of Indications of Whole-Body MRI Beyond Oncology: Imaging Answers to Clinical Needs. Semin Musculoskelet Radiol 19:348–362
12. Poggenborg RP, Wiell C, Boyesen P et al (2014) No overall damage progression despite persistent inflammation in adalimumab-treated psoriatic arthritis patients: results from an investigator-initiated 48-week comparative magnetic resonance imaging, computed tomography and radiography trial. Rheumatology (Oxford) 53:746–756
13. Bozgeyik Z, Ozgocmen S, Kocakoc E (2008) Role of diffusion-weighted MRI in the detection of early active sacroiliitis. AJR Am J Roentgenol 191:980–986
14. Gezmis EDF, Agildere M (2013) Diagnosis of early sacroiliitis in seronegative spondyloarthropathies by DWI and correlation of clinical and laboratory findings with ADC values. Eur J Radiol 82:2316–2321
15. Sanal HT, Yilmaz S, Kalyoncu U et al (2013) Value of DWI in visual assessment of activity of sacroiliitis in longstanding ankylosing spondylitis patients. Skeletal Radiol 42:289–293
16. Sanal HT, Yilmaz S, Simsek I et al (2013) Apparent diffusion coefficients of sacroiliitis in patients with established ankylosing spondylitis. Clin Imaging 37:734–739
17. Vendhan K, Bray TJ, Atkinson D et al (2016) A diffusion-based quantification technique for assessment of sacroiliitis in adolescents with enthesitis-related arthritis. Br J Radiol 89:20150775
18. Padhani AR, Koh DM, Collins DJ (2011) Whole-body diffusion-weighted MR imaging in cancer: current status and research directions. Radiology 261:700–718
19. Lecouvet FE (2016) Whole-Body MR Imaging: Musculoskeletal Applications. Radiology 279:345–365
20. Eshed I, Hermann KG (2015) Novel imaging modalities in spondyloarthritis. Curr Opin Rheumatol 27:333–342
21. van der Linden S, Valkenburg HA, Cats A (1984) Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 27:361–368
22. Machado P, Landewe R, Lie E et al (2011) Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 70:47–53
23. Rudwaleit M, Schwarzlose S, Hilgert ES, Listing J, Braun J, Sieper J (2008) MRI in predicting a major clinical response to anti-tumour necrosis factor treatment in ankylosing spondylitis. Ann Rheum Dis 67:1276–1281
24. Landewe RB, Hermann KG, van der Heijde DM et al (2005) Scoring sacroiliac joints by magnetic resonance imaging. A multiple-reader reliability experiment. J Rheumatol 32:2050–2055
25. Mann HBWDR (1947) On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. Ann Math Stat 18:50–60
26. Wilcoxon F (1946) Individual comparisons of grouped data by ranking methods. J Econ Entomol 39:269

27. Kwee TC (2014) Can whole-body MRI replace (18)F-fluorodeoxyglucose PET/CT? Lancet Oncol 15:243–244
28. Rudwaleit M, Jurik AG, Hermann KG et al (2009) Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 68:1520–1527
29. Lambert RG, Bakker PA, van der Heijde D et al (2016) Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 75:1958–1963
30. Sahin N, Hacibeyoglu H, Ince O et al (2015) Is there a role for DWI in the diagnosis of sacroiliitis based on ASAS criteria? Int J Clin Exp Med 8:7544–7552
31. Poggenborg RP, Eshed I, Ostergaard M et al (2015) Enthesitis in patients with psoriatic arthritis, axial spondyloarthritis and healthy subjects assessed by 'head-to-toe' whole-body MRI and clinical examination. Ann Rheum Dis 74:823–829