RMD Open
Rheumatic & Musculoskeletal Diseases

ORIGINAL ARTICLE

What is the reliability of non-trained investigators in recognising structural MRI lesions of sacroiliac joints in patients with recent inflammatory back pain? Results of the DESIR cohort

Charlotte Jacquemin,[1] Roxana Rubio Vargas,[2] Rosaline van den Berg,[2] Fabrice Thévenin,[3] Gregory Lenczner,[4] Monique Reijnierse,[5] Salah Ferkal,[6] Philippe Le Corvoisier,[6] Alain Rahmouni,[7] Damien Loeuille,[8] Antoine Feydy,[3] Maxime Dougados,[9] Désirée van der Heijde,[2] Pascal Claudepierre[1,10]

To cite: Jacquemin C, Rubio Vargas R, van den Berg R, *et al*. What is the reliability of non-trained investigators in recognising structural MRI lesions of sacroiliac joints in patients with recent inflammatory back pain? Results of the DESIR cohort. *RMD Open* 2016;**2**:e000303. doi:10.1136/rmdopen-2016-000303

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/rmdopen-2016-000303).

CJ and RRV contributed equally to this study.

Received 25 April 2016
Revised 8 September 2016
Accepted 12 September 2016

For numbered affiliations see end of article.

Correspondence to
Professor Pascal Claudepierre;
pascal.claudepierre@aphp.fr

ABSTRACT

Objective: The objective of this study was to evaluate the reliability of recognising structural lesions on MRI (erosions, fatty lesions, ankylosis) of the sacroiliac joints (MRI-SIJ) in clinical practice compared to a central reading in patients with a possible recent axial spondyloarthritis (axSpA).

Methods: Patients aged 18–50 years, with recent (<3 years) and chronic (≥3 months) inflammatory back pain, suggestive of axSpA were included in the DEvenir des Spondyloarthrites Indifférenciées Récentes (DESIR) cohort. MRI-SIJ structural lesions were scored by non-trained local readers, and by two trained central readers. Local readers scored each SIJ as normal, doubtful or definite lesions. Central readers scored separately each type of lesion. The central reading (mean of the two central readers' scores) was the external standard. Agreement (κ) was calculated first between local (3 definitions of a positive MRI-SIJ) and central readings (9 definitions), and then between the two central readers.

Results: 664/708 patients with complete available images were included. Agreements between local and central readings were overall 'fair', except when considering at least 2 or 3 fatty lesions and at least 3 erosions and/or fatty lesions where agreement was 'moderate'. Agreement between central readers was similar. MRI-SIJ was positive for 52.6% of patients according to central reading (at least 1 structural lesion) and for 35.4% of patients according to local reading (at least unilateral 'doubtful' or 'definite' structural lesions).

Conclusions: Agreement on a positive structural MRI-SIJ was fair to moderate between local and central readings, as well as between central readers. The reliability improved when fatty lesions were considered.

Trial registration number: NCTO 164 8907.

Key messages

What is already known about this subject?

- Recent data suggest that structural lesions on MRI of the sacroiliac joints (MRI-SIJ) may be used instead of structural damage on conventional radiographs in the classification of patients.[1]

What does this study add?

- Reliability of non-trained readers as well as trained readers is fair to moderate to recognise structural lesions of SIJ on MRI in patients with recent inflammatory back pain suggestive of spondyloarthritis.
- Reliability between local and central readings and between central readers is better when considering at least two fatty lesions.

How might this impact on clinical practice?

- Structural lesions are difficult to identify on MRI-SIJ and need to be better defined to be used for axial spondyloarthritis diagnosis in clinical practice.

INTRODUCTION

Diagnosing axial spondyloarthritis (axSpA) in an early stage of the disease is a challenge in clinical practice. Evolution of axSpA can lead to irreversible structural damage and impact the quality of life. There is more and more evidence that early treatment may change the outcome in patients with axSpA.[2 3] The Assessment of SpondyloArthritis international Society (ASAS) developed classification criteria relevant for early axSpA.[4] According to these criteria, patients with at least one SpA

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

feature may be classified in case of structural lesions on X-ray of the sacroiliac joints (X-SIJ) or active sacroiliitis on MRI.[5] Although only inflammatory lesions (bone marrow oedema (BME) and osteitis) have been selected in the ASAS criteria, structural lesions (erosions, fatty lesions, sclerosis, ankylosis) are also visible on MRI of the sacroiliac joints (MRI-SIJ), and have been described by the ASAS/OMERACT MRI group.[6] The European League Against Rheumatism recommends taking into account inflammatory lesions as well as structural lesions when diagnosing axSpA.[7] Radiographic structural lesions of SIJ are still assessed by the same grading, first described in the New York criteria.[8] However, it has been shown to be difficult to identify and discriminate structural lesions on X-SIJ between different observers,[9] and training does not seem to improve it.[10] These findings were recently confirmed in patients with recent-onset chronic inflammatory back pain (IBP) of the DEvenir des Spondyloarthrites Indifferenciées Récentes; Outcomes in patients with Recent-onset Undifferentiated Spondyloarthritis (DESIR) cohort.[11] Furthermore, radiographic sacroiliitis may appear several years after the onset of symptoms,[12] delaying the diagnosis of ankylosing spondylitis (AS).[13] MRI-SIJ could be an alternative to X-SIJ, especially in early axSpA. It has been recently demonstrated in the DESIR cohort that MRI-SIJ could be reliably used instead of, or in addition to, X-SIJ to assess structural lesions without significant changes in the classification according to ASAS axSpA criteria.[1] Previous studies have shown a good reliability of MRI-SIJ to detect structural lesions.[14–16] However, in these previous studies, MRI-SIJ were scored by trained readers, blinded for clinical information, while in daily practice scoring is performed by local radiologists or rheumatologists, who have access to clinical data. The ability of these local non-trained readers (ie, not specifically trained for the assessment of structural lesions on MRI-SIJ, but trained as a radiologist or rheumatologist) in various centres to detect structural lesions on MRI-SIJ, instead of a centralised reading, is unknown. The main objective of this study was to evaluate the reliability of non-trained investigators in recognising structural lesions on MRI-SIJ of patients with recent inflammatory back pain, suggestive of axSpA in the DESIR cohort, compared to trained central readers. To put the data in context, the agreements between the two central readers are also presented as well as agreements between 'MRI-SIJ structural lesions', according to the local reading, and 'radiographic sacroiliitis' according to the local and central reading.

METHODS

Study design

This study is a cross-sectional study of baseline data of the multicentre prospective longitudinal DESIR cohort.[17]

Patients

All patients of the DESIR cohort with available MRI-SIJ and X-SIJ at baseline were included. The DESIR cohort has already been described extensively.[17] In short, patients aged between 18 and 50 years and suffering from recent (<3 years) and chronic (>3 months) IBP (in the buttocks, lumbar or thoracic spine) fulfilling either the Calin[18] or the Berlin criteria[19] were recruited in 25 regional centres in France. Symptoms must be suggestive of axSpA according to the local investigator's assessment with a score ≥5/10 (0=not suggestive; 10=very suggestive for axSpA).[17] In total, 708 patients were included between December 2007 and April 2010. The baseline database was locked on 30 October 2012. The cohort was approved by an ethics committee and complied with good clinical practices. A detailed description of the organisation of the cohort, centres, protocol and collected data is available at http://www.lacohortedesir.fr/desir-in-english/. This study was approved by the scientific committee of the DESIR cohort. All patients signed informed consent.

Collected data

Baseline demographic parameters including age and gender, clinical parameters and laboratory test results including acute phase reactants and human leucocyte antigen (HLA)-B27 antigen status were collected using a standardised case report form.

MRI-SIJ assessment

MRI-SIJ assessment were performed on a 1 or 1.5 T MRI machine, using coronal oblique T1-weighted fast spin echo and short τ inversion recovery (STIR) sequences, with 12–15 slices of 4 mm thickness. MRI-SIJ structural lesions were scored first by the local reader and subsequently by two central readers.

The local readers were radiologists or rheumatologists in each centre with possible access to clinical data. They did not participate in any training session and were instructed to score each SIJ in 3 grades, according to the presence/absence of structural lesions (defined as typical sclerosis, erosions, bony bridges or ankylosis). Grade 0 was corresponding to normal, grade 1 to doubtful and grade 2 to definite structural lesions.

The central readers first participated in a calibration session (see below). The scoring was mainly based on a T1-weighted sequence, but the readers had free access to the STIR sequence. The central readers were blinded for clinical and other imaging data and for the local readers' results. We have used the adapted SpondyloArthritis Research Consortium of Canada scoring system by Weber *et al*[20] to assess MRI-SIJ. This method is based on the assessment of lesions (present vs absent) on six consecutive slices through the SIJ, in its cartilaginous compartment (ie, its anteroinferior portion), starting on the slice on which at least 1 cm of the cartilage is visible, from anterior to posterior. Each SIJ is divided into four quadrants. Fatty lesions were only marked present by the readers if appearing with a distinct border and a homogeneous pattern. The readers only marked fatty lesions and erosions present if

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

observed on at least two consecutive slices, resulting in a maximum score of 40 per lesion per patient (20 per SIJ). Ankylosis was considered present if seen on a single slice, with a maximum score of 24 per patient (note that ankylosis always involves two quadrants, either upper iliac and upper sacral or lower iliac and lower sacral quadrants). The presence of sclerosis was not taken into account for this study, since the interest of including sclerosis in structural scores has been debated: most scores did not include it[14 20–22] and when it has been included, sclerosis usually showed low reliability.[23] The mean score of the two central readers was used for each lesion to obtain the central reading score; no adjudicator has been used.

Several definitions were used to define a positive MRI-SIJ for structural lesions by local and central readers, at both the SIJ and patient levels (table 1).

Definition of a positive MRI-SIJ for structural lesions

Concerning the local reading, the MRI-SIJ was considered positive if at least one of the two SIJs was scored 'doubtful or abnormal', that is, 'at least unilateral grade 1' (definition 1). In sensitivity analyses, we also considered: 'at least unilateral grade 2' (definition 2), and 'bilateral grade 2' (definition 3) as a positive MRI-SIJ (table 1).

For the central reading, in the absence of a generally accepted definition of a positive structural MRI-SIJ, we considered the following definition: 'presence of ≥1 erosion or ≥1 fatty lesion or ≥1 ankylosis in at least one SIJ' (definition 1). We also considered eight additional definitions (table 1).

X-SIJ assessment

X-SIJs were performed in the anteroposterior view. Similar to MRI-SIJ, baseline X-SIJ were read by the local investigators, and later by two trained central readers. The procedure has been previously described in detail.[11] In short, the local readers scored each SIJ according to a method derived from the modified New York criteria (mNY):[24] grades 2 and 3 were pooled in one combined grade 2. Thus, grade 0 was corresponding to normal, grade 1 to doubtful, grade 2 to definite sacroiliitis and grade 3 to ankylosis. The central readers used the mNY criteria scoring method, in which sacroiliitis is defined as grade ≥2 bilaterally or grades 3–4 unilaterally. In case of disagreement, an experienced radiologist served as the adjudicator. We considered several definitions of a positive X-SIJ for local and central readings (table 1).

Definition of a patient with radiographic sacroiliitis

For the local reading, radiographic sacroiliitis was defined as at least one SIJ scored 'grade 2 DESIR' (definition 1). As sensitivity analyses, we also considered two other definitions (table 1). For the central reading, sacroiliitis was defined according to the mNY criteria (table 1).

Training of the central readers

The central readers of the MRI-SIJ and X-SIJ have participated in calibration sessions. The calibration process with the same readers has been extensively described previously.[11 25] After agreements between central readers were moderate to good (κ=0.4–1.0) regarding the various types of structural lesions on MRI-SIJ, the two central readers could start reading all available baseline MRI-SIJ of the DESIR cohort.

Statistical analysis

Inter-rater agreements were calculated using Cohen's κ,[26 27] and positive and negative percent agreements

Table 1 Definitions of positive MRI-SIJ and X-SIJ for structural lesions in local and central readings

Definitions of positive MRI-SIJ Patient level		Definitions of positive X-SIJ	
Local reading		*Local reading*	
Definition 1	At least unilateral grade ≥1	Definition 1	At least unilateral 'DESIR grade' 2
Definition 2	At least unilateral grade 2	Definition 2	At least bilateral 'DESIR grade' 2 or unilateral 'DESIR grade' 3
Definition 3	Bilateral grade 2	Definition 3	Bilateral 'DESIR grade' 3
Central reading		*Central reading*	
Definition 1	Presence of ≥1 erosion or fatty lesion or ankylosis (partial or total)	Radiographic sacroiliitis	Bilateral 'mNY grade ≥2' or at least unilateral 'mNY grade≥3'
Definition 2	≥1 erosion		
Definition 3	≥2 erosions		
Definition 4	≥3 erosions		
Definition 5	≥1 fatty lesion		
Definition 6	≥2 fatty lesions		
Definition 7	≥3 fatty lesions		
Definition 8	Presence of ankylosis (partial or total)		
Definition 9	≥3 erosions and/or fatty lesions		

mNY, modified New York criteria.[24] MRI-SIJ, MRI of the sacroiliac joints.

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

(PPA/NPA).[28] All κs were interpreted according to the standards proposed by Landis and Koch; values<0 indicated no agreement, 0–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial and 0.81–1 almost perfect agreement.[27]

Central reading was assessed by the mean score of central readers. For sensitivity analyses, we also considered lesions recognised by both central readers and lesions scored by only central reader 1 or only central reader 2.

For the primary objective, agreement was calculated between the local reading and the central reading, which served as an external standard, regarding the different definitions of a positive MRI-SIJ. Next, we investigated which type of lesions (reported by the central reading) contributed most to the global assessment of the local readers (3 definitions). As a sensitivity analysis, agreement at the SIJ level was also assessed. Then we calculated the agreement between central readers regarding the nine definitions of a positive structural MRI-SIJ with regard to the number of each structural lesion (erosion, fatty lesions and ankylosis) with various cut-off values (≥1 to ≥5) based on a patient's level. Prevalence of each type of lesion (score≥1) was also assessed.

Finally, agreement was calculated between the various definitions of a positive MRI-SIJ and a positive X-SIJ, first for the local reading and second for the central reading as the external standard. Sensitivity and specificity of MRI-SIJ compared to X-SIJ were calculated, with the X-SIJ as the external standard.

Analyses were performed using Stata SE V.12.

RESULTS

Complete MRI-SIJ baseline data were available in 664 patients. Their mean age was 34.9 (SD 8.7) years and mean symptom duration 24 (SD 22.3) months; 309 (46.5%) patients were men and 388 (51.4%) were HLA-B27 positive; 409/582 (70.2%) were fulfilling the ASAS criteria for axSpA.[25]

Prevalence of structural lesions scored by central readers

On the basis of the mean scores of the two central readers, 52.6% of patients had at least one structural lesion including erosions (32.2%), fatty lesions (29.2%) and ankylosis (24.2%) (table 2). When looking at the lesions agreed on by both central readers, 29.1% of patients had at least one structural lesion including erosions (9.3%), fatty lesions (14.6%) and ankylosis (5.2%). For each type of lesion, central reader 1 scored systematically fewer lesions than central reader 2.

Agreement between local reading and central reading in recognising structural lesions on MRI-SIJ.

The κ coefficients, PPA and NPA, between the local reading and the central reading for each definition are shown in table 3 (complete data in supplementary table 1). Local and central readings were concordant for 66.3% of patients according to the presence/absence of abnormalities (definitions 1). Overall, agreements were better for definitions considering at least two or at least three fatty lesions and at least three erosions and/or fatty lesions in central reading and if a more stringent definition was considered in local reading. According to the local reading (definition 1), 35.4% of patients had a positive MRI-SIJ for structural lesions and according to the central reading (definition 1), 52.6% of patients had a positive MRI-SIJ for structural lesions; 8.3% of patients with a positive MRI-SIJ according to the local reading were not recognised by the central reading and 25.5% of patients with a negative MRI-SIJ according to the local reading were marked positive by the central reading.

Agreement between the two central readers for various definitions of a positive MRI-SIJ for structural lesions

Agreement between the two central readers was 'slight' to 'fair' (κ=0.19–0.29) when considering definitions including at least one structural lesion (fatty lesion, erosion or ankylosis) (definition 1) or only erosions (definitions 2, 3 and 4; table 4). For definitions considering at least two or at least three fatty lesions and at least three erosions and/or fatty lesions, agreement was better but still only moderate (κ=0.44–0.51). With an increasing number of fatty lesions (from ≥1 to ≥5 fatty lesions), agreement increased (κ=0.40–0.59) (supplementary table 4). Conversely, with an increasing number of erosions (from ≥1 to ≥5 erosions), the agreement decreased (κ=0.22–0.07) (supplementary table 5). When considering definition eight including only ankylosis, agreement was 'slight'.

Sensitivity analyses

At the SIJ level, agreements between local and central readings and between the two central readers were similar between the left and right SIJ (data not shown).

Table 2 Prevalence of various MRI structural lesions scored by central readers at the patient level

Patients with at least one of the following types of lesions	Reader 1, N (%)	Reader 2, N (%)	Agreement of both central readers*, N (%)	Mean of the two central readings†, N (%)
Any type of lesions	277 (41.7)	348 (52.4)	193 (29.1)	349 (52.6)
Erosions	82 (12.3)	256 (38.6)	62 (9.3)	214 (32.2)
Fatty lesions	162 (24.4)	186 (28)	97 (14.6)	194 (29.2)
Ankylosis (partial or total)	122 (18.4)	172 (25.9)	39 (5.9)	161 (24.2)

*Number of patients with at least one lesion scored by both readers.
†Number of patients with at least one lesion based on the mean score of the two central readers.

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

Table 3 Agreements between local and central readings for various definitions of a positive MRI-SIJ for structural lesions

	MRI central reading (mean of the two central readers)					
	Definition 1: ≥1 erosion or ≥1 fatty lesion or ankylosis		**Definition 2: ≥1 erosion**		**Definition 3: ≥2 erosions**	
MRI local reading	**PPA (%)**	**NPA (%)**	**PPA (%)**	**NPA (%)**	**PPA (%)**	**NPA (%)**
Definition 1: at least unilateral grade ≥1	61.6	69.9	54.3	76.7	48.6	80.5
	κ (95% CI): 0.34 (0.27 to 0.40)		κ (95% CI): 0.31 (0.24 to 0.39)		κ (95% CI): 0.31 (0.24 to 0.39)	
Definition 2: at least unilateral grade 2	53.5	69.8	54.3	79.9	49.8	85.0
	κ (95% CI): 0.29 (0.22 to 0.35)		κ (95% CI): 0.35 (0.27 to 0.43)		κ (95% CI): 0.35 (0.27 to 0.43)	
Definition 3: bilateral grade 2	44.3	69.5	48.0	80.8	50.0	88.3
	κ (95% CI): 0.24 (0.18 to 0.29)		κ (95% CI): 0.31 (0.24 to 0.39)		κ (95% CI): 0.38 (0.30 to 0.47)	
	Definition 4: ≥3 erosions		**Definition 5: ≥1 fatty lesion**		**Definition 6: ≥2 fatty lesions**	
Definition 1	38.0	80.2	55.9	79.0	52.3	81.8
	κ (95% CI): 0.24 (0.17 to 0.30)		κ (95% CI): 0.35 (0.28 to 0.43)		κ (95% CI): 0.36 (0.29 to 0.43)	
Definition 2	40.8	85.6	56.5	83.3	54.2	86.3
	κ (95% CI): 0.28 (0.20 to 0.37)		κ (95% CI): 0.40 (0.32 to 0.48)		κ (95% CI): **0.41** (0.33 to 0.49)	
Definition 3	46.4	90.1	52.1	85.1	56.9	89.9
	κ (95% CI): 0.37 (0.28 to 0.46)		κ (95% CI): 0.38 (0.30 to 0.46)		κ (95% CI): **0.47** (0.38 to 0.55)	
	Definition 7: ≥3 fatty lesions		**Definition 8: ankylosis**		**Definition 9: ≥3 erosions and/or fatty lesions**	
Definition 1	46.9	82.9	30.3	70.4	56.8	81.5
	κ (95% CI): 0.34 (0.27 to 0.41)		κ (95% CI): 0.02 (-0.05 to 0.10)		κ (95% CI): 0.39 (0.32 to 0.47)	
Definition 2	50.2	87.7	26.3	75.1	57.6	85.6
	κ (95% CI): 0.40 (0.31 to 0.48)		κ (95% CI): 0.01 (-0.06 to 0.09)		κ (95% CI): **0.43 (0.35 to 0.51)**	
Definition 3	56.2	91.8	24.1	79.5	57.0	88.3
	κ (95% CI): **0.48 (0.39 to 0.57)**		κ (95% CI): 0.04 (-0.04 to 0.12)		κ (95% CI): **0.46 (0.38 to 0.54)**	

κ in bold represent κ >0.40.
MRI-SIJ, MRI of the sacroiliac joints; NPA, negative percent agreement; PPA, positive percent agreement.

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

Table 4 Agreement between central readers for various definitions of a positive MRI-SIJ for structural lesions on patient level

	Definition 1: ≥1 erosion or ≥1 fatty lesion or ankyloses	Definition 2: ≥1 erosion	Definition 3: ≥2 erosions	Definition 4: ≥3 erosions	Definition 5: ≥1 fatty lesion
PPA (%)	61.8	36.7	32.7	25.7	55.7
NPA (%)	66.0	78.4	84.3	88.4	84.3
κ (95% CI)	0.29 (0.22 to 0.36)	0.22 (0.16 to 0.29)	0.23 (0.15 to 0.30)	0.19 (0.11 to 0.27)	0.40 (0.32 to 0.48)
	Definition 6: ≥2 fatty lesions	**Definition 7: ≥3 fatty lesions**	**Definition 8: ankylosis**	**Definition 9: ≥3 erosions and/or fatty lesions**	
PPA (%)	55.3	58.8	26.5	56.8	
NPA (%)	88.9	92.2	79.1	85.7	
κ (95% CI)	**0.45** (0.36 to 0.53)	**0.51** (0.42 to 0.60)	0.06 (−0.01 to 0.14)	**0.44** (0.36 to 0.51)	

κ in bold represent κ >0.40.
MRI-SIJ, MRI of the sacroiliac joints; NPA,negative percent agreement; PPA, positive percent agreement.

Table 5 Agreement between various definitions of a positive MRI-SIJ for structural lesions (local reading) and radiographic sacroiliitis

	X-ray local reading						X-ray central reading (2/3)	
	Definition 1: at least unilateral 'DESIR grade' 2		Definition 2: at least bilateral 'DESIR grade' 2 or unilateral 'DESIR grade' 3		Definition 3: bilateral 'DESIR grade' 3		bilateral 'mNY grade ≥2' or at least unilateral 'mNY grade≥3'	
MRI local reading	**PPA (%)**	**NPA (%)**	**PPA (%)**	**NPA (%)**	**PPA (%)**	**NPA (%)**	**PPA (%)**	**NPA (%)**
Definition 1: at least unilateral grade ≥1	63.9	83.5	51.9	83.3	2.6	78.5	56.0	82.6
	κ (95% CI): **0.48** (0.41 to 0.55)		κ (95% CI): 0.38 (0.31 to 0.45)		κ (95% CI): 0.01 (−0.01 to 0.03)		κ (95% CI): 0.40 (0.33 to 0.47)	
	Se=73.9%. Sp=78.6%		Se=83.7%. Sp=73.6%		Se=75.0%. Sp=64.7%		Se=75.2%. Sp=75.0%	
Definition 2: at least unilateral grade 2	66.9	87.9	56.7	88.4	3.4	85.0	60.4	87.7
	κ (95% CI): **0.55** (0.48 to 0.62)		κ (95% CI):**0.46** (0.38 to 0.54)		κ (95% CI): 0.02 (−0.01 to 0.05)		κ (95% CI): **0.48** (0.40 to 0.56)	
	Se=65.9%. Sp=88.4%		Se=75.0%. Sp=82.9%		Se=75.0%. Sp=74.0%		Se=67.9%. Sp=84.8%	
Definition 3: bilateral grade 2	59.0	87.9	61.9	92.1	4.9	90.0	60.9	90.4
	κ (95% CI): **0.48** (0.40 to 0.55)		κ (95% CI):**0.54** (0.45 to 0.63)		κ (95% CI): 0.04 (−0.01 to 0.08)		κ (95% CI): **0.51** (0.40 to 0.60)	
	Se=49.4%. Sp=93.2%		Se=66.3%. Sp=90.8%		Se=75.0%. Sp=82.0%		Se=56.9%. Sp=92.0%	

κ in bold represent κ >0.40.
MRI-SIJ, MRI of the sacroiliac joints; NPA, negative percent agreement; PPA, positive percent agreement; Se, sensitivity; Sp, specificity.

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

Comparing local reading and each central reader separately, central reader 1 had lower agreement for definitions 3 and 4 (≥2 or ≥3 erosions) than the mean of both central readers (supplementary tables 2 and 3).

Agreement between a positive MRI-SIJ for structural lesions by local Reading and radiographic sacroiliitis

In total, 649 patients had complete data for this analysis (ie, local reading of MRI-SIJ and local and central readings of X-SIJ). About 21.1% of patients had AS according to mNY criteria in the central reading. Compared to the radiographic local reading, agreements were 'moderate' at best (from 0.38 to 0.55) when considering the two less stringent radiographic definitions (definitions 1 and 2) (table 5). Compared to the radiographic central reading, at the patient level, κs were moderate for the various MRI-SIJ definitions (from 0.40 to 0.51). In both comparisons, MRI-SIJ sensitivity was better for a less stringent MRI-SIJ definition and specificity was better for a more stringent MRI-SIJ definition.

DISCUSSION

In patients with recent chronic IBP suggestive of axSpA, the agreement between readings of structural lesions on MRI-SIJ by two trained central readers and by local readers is 'moderate' at best. The agreement between local readers and the various definitions of the central reading were always low ('slight' or 'fair') except for the definitions considering '2' or '3' fatty lesions or 'at least 3 erosions and/or fatty lesions', where the agreements reached a 'moderate' level. These results show that the reliability of the routine reading, in this context, is overall not good. Although local readers did not independently score each type of lesion, these results also suggest that local readers may recognise fatty lesions better than other types of structural lesions. However, we have to keep in mind that local and central scorings were different, which may have impacted the assessment of agreement.

Agreement between central readers is overall not very different from agreement between local and central readings. Surprisingly, a 'fair' agreement between the two central readers was found for the first definition, although it was the loosest definition ('presence of at least one erosion or at least one fat deposition, or presence of partial or total ankylosis') and the lowest agreement was found for definition 8 regarding the presence of ankylosis. Despite the calibration sessions of the readers, reader 1 systematically scored fewer lesions than reader 2, in particular with regard to erosions. Thus, reader 1 was more specific and reader 2 more sensitive, leading to a low agreement between the two readers. Reader 2 may have considered physiological abnormalities as structural lesions (especially for erosions), which might question the validity of the calibration process. However, the same two central readers also scored BME lesions in the DESIR cohort with a similar calibration method and obtained a much better κ (0.73).[25] Moreover, despite this limitation, considering the agreements between local reading and each central reader separately, the agreement for erosions was better with reader 2 than with reader 1 and agreement for fatty lesions was similar. This suggests that the fairly low agreement may be rather due to the general difficulty in distinguishing structural lesions from physiological abnormalities and not (only) to the calibration and reading methods. Moreover, global agreement for structural lesions (including erosions, fatty lesions and ankylosis) seems better in studies including a high proportion of patients with AS (κ=0.84),[14] or when the median symptoms duration is longer (κ=0.76–0.80).[29] Puhakka *et al*[23] included patients with a short symptom duration (median disease duration=19 months) excluding patients with AS, and found an agreement similar to our study (κ=0.42).

Contrary to erosions and ankylosis, fatty lesions contributed most to a better agreement between central readers and between central and local MRI readings. The consideration of at least two or, even better, three fatty lesions increased the reliability between readers. Those results are in agreement with the findings of a recent study[30] which defined a cut-off of at least three fatty lesions or 5 fatty lesions and/or erosions to define a positive MRI-SIJ with <5% of patients without axSpA fulfilling this definition. In previous studies, agreement seemed usually better for erosions in AS and for fatty lesions in recent axSpA.[14 16 23 31 32] Weber *et al*[33] suggested that assessment of fatty lesions in the SIJ may have a diagnostic utility in early axSpA with doubtful erosions or BME, but that only fatty lesions with a distinct border or a homogeneous pattern (and not subchondral lesions) should be considered. In our study, we only marked fatty lesions if there were a distinct border and a homogeneous pattern. Recently, a T1-weighted opposed-phase gradient-echo (opGE) sequence has shown promising results and may improve reliability in detecting erosions.[34]

Prevalence of structural lesions was low in our study compared to most of the previously published studies. This difference can be partly explained by the lower proportion of patients with AS in our study (21.1% according to mNY criteria by central readers, and 70.2% of axSpA according to the ASAS criteria[25]), whereas other studies included 60.7–73.5% of patients with AS.[15 32 35] As expected, the prevalence of structural lesions is higher in patients with AS than in non-radiographic patients with axSPA.[29 30 32] Part of these results can be explained by the pathophysiology of axSpA: fatty lesions appear to follow resolution of inflammation and of erosions in axSpA.[36 37]

Previous studies have compared MRI central readings to X-ray central readings with heterogeneous results (Se=49–84% and Sp=61–98%).[12 19] Agreements between MRI-SIJ local readings and X-SIJ local readings and between MRI-SIJ local readings and X-SIJ central readings in our study were moderate. These results confirm

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

that structural lesions are difficult to assess by MRI in clinical practice (local reading) when compared to the gold standard (X-SIJ central reading) and to what is currently performed in clinical practice (local X-SIJ reading). However, the low reliability of X-SIJ readings, which has been confirmed in the DESIR cohort,[11] may interfere with the current analysis. Furthermore, this analysis was limited by the use of non-validated definitions for the X-SIJ local reading.

A limitation of this study was the absence of comparison with CT. CT is known to be reliable for the assessment of the bony structures, including erosions, sclerosis and ankylosis and is a better external standard than radiographs.[38 39] However, it has been chosen in the DESIR cohort not to add CT scan because of time, costs and irradiation dose.

It should also be noted that both local and central readers had access to the STIR sequences and that the presence of BME lesions may have biased their scoring by influencing the global appraisal. However, this is also the situation in which the MRI-SIJ will be used in clinical practice. Moreover, the local readers also might have had access to the radiographs while scoring the MRI-SIJ, while the central readers were completely blinded for the other imaging modality.

Finally, this study may have been limited by the restriction of structural lesions scoring to their presence on two consecutive slices (which is the rule for BME lesions), contrary to previous studies in which their presence on a single slice was often sufficient. This may have hampered the comparison with previous studies.

This study also has several strengths. The DESIR cohort has included a large number of patients in whom a standardised MRI-SIJ was systematically performed at baseline. Previous studies reporting on this topic mostly included <100 patients with axSpA.[14 16 23 29 31 40] The DESIR patients had recent disease onset with symptoms suggestive of axSpA at baseline, but the diagnosis was not necessarily confirmed. The assessment of structural lesions on MRI-SIJ is of particular interest in this type of patients to assist in diagnosing axSpA. Results from previous studies assessing structural lesions in patients with a confirmed diagnosis and long-standing disease[14 29 31] cannot be extrapolated to this category of patients.

In conclusion, the reliability of non-trained readers as well as trained readers to recognise structural lesions on MRI-SIJ in patients with recent IBP suggestive of SpA is overall 'fair' to 'moderate'. Fatty lesions may contribute most to a positive structural MRI-SIJ assessment by non-trained and trained readers.

Author affiliations
[1]Department of Rheumatology, Hôpital Henri Mondor, Créteil, France
[2]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[3]Department of Radiology, Hôpital Cochin, Paris, France
[4]Department of Radiology, Clinique Hartmann, Neuilly-sur-Seine, France
[5]Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands
[6]Clinical Investigation Center 1430, AP-HP and Inserm, Hôpital Henri Mondor, Créteil, France
[7]Department of Radiology, Hôpital Henri Mondor, Créteil, France
[8]Department of Rheumatology, Hôpital Brabois, Nancy, France
[9]Department of Rheumatology, Hôpital Cochin, Paris, France
[10]Université Paris Est Créteil, EA 7379—EpidermE, F-94010, Créteil, France

Twitter Follow Damien Loeuille at @no

Acknowledgements The DESIR cohort is conducted under the control of Assistance Publique–Hôpitaux de Paris via the Clinical Research Unit Paris-Centre and under the umbrella of the French Society of Rheumatology and INSERM (Institut National de la Santé et de la Recherche Médicale). The database management is performed within the department of epidemiology and biostatistics (Professor Jean-Pierre Daurès, D.I.M., Nîmes, France). The authors wish to thank the 25 participating centres and their investigators: Pr Maxime Dougados (Paris—Cochin B), Pr André Kahan (Paris—Cochin A), Pr Olivier Meyer (Paris—Bichat), Pr Pierre Bourgeois (Paris—La Pitié-Salpetrière), Pr Francis Berenbaum (Paris—Saint Antoine), Pr Pascal Claudepierre (Créteil), Pr Maxime Breban (Boulogne-Billancourt), Dr Bernadette Saint-Marcoux (Aulnay-sous-Bois), Pr Philippe Goupille (Tours), Pr Jean-Francis Maillefert (Dijon), Dr Xavier Puéchal (Le Mans), Pr Daniel Wendling (Besançon), Pr Bernard Combe (Montpellier), Pr Liana Euller-Ziegler (Nice), Pr Philippe Orcel (Paris—Lariboisière), Pr Pierre Lafforgue (Marseille), Dr Patrick Boumier (Amiens), Pr Jean-Michel Ristori (Clermont-Ferrand), Dr Nadia Mehsen (Bordeaux), Pr Damien Loeuille (Nancy), Pr René-Marc Flipo (Lille), Pr Alain Saraux (Brest), Pr Corinne Miceli (Le Kremlin Bicêtre), Pr Alain Cantagrel (Toulouse), Pr Olivier Vittecoq (Rouen).

Contributors PCP drafted the study design. CJ drafted the manuscript with important contributions from PCP, DvdH, RvdB and RRV. RRV and RvdB analysed and interpreted the data. RvdB, GL, FT and MR read the X-ray and/or MRI. AR, DL, AF, MD and PCP participated in the data collection. All authors approved the final manuscript.

Funding This specific work has not been funded, but the DESIR-cohort is financially supported by unrestricted grants from both the French Society of Rheumatology and Pfizer France. Neither funding source had any role in designing the study; collecting, analysing or interpreting the data; writing the manuscript; or deciding to submit the manuscript for publication.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES

1. Bakker PA, van den Berg R, Lenczner G, *et al*. Can we use structural lesions seen on MRI of the sacroiliac joints reliably for the classification of patients according to the ASAS axial spondyloarthritis criteria? Data from the DESIR cohort. *Ann Rheum Dis* Published Online First: 4 Aug 2016. doi: 10.1136/annrheumdis-2016-209405.
2. Robinson PC, Brown MA. The window of opportunity: a relevant concept for axial spondyloarthritis. *Arthritis Res Ther* 2014;16:109.
3. Seo MR, Baek HL, Yoon HH, *et al*. Delayed diagnosis is linked to worse outcomes and unfavourable treatment responses in patients with axial spondyloarthritis. *Clin Rheumatol* 2015;34:1397–405.
4. Rudwaleit M, van der Heijde D, Landewé R, *et al*. The Assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. *Ann Rheum Dis* 2011;70:25–31.
5. Rudwaleit M, van der Heijde D, Landewé R, *et al*. The development of Assessment of SpondyloArthritis International Society classification criteria for axial spondyloarthritis (part II): validation and final selection. *Ann Rheum Dis* 2009;68:777–83.

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.

6. Rudwaleit M, Jurik AG, Hermann KGA, *et al*. 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* 2009;68:1520–7.
7. Mandl P, Navarro-Compán V, Terslev L, *et al*. EULAR recommendations for the use of imaging in the diagnosis and management of spondyloarthritis in clinical practice. *Ann Rheum Dis* 2015;74:1327–39.
8. Bennett P, Wood P. *Population studies of the rheumatic diseases*. Amsterdam: Excerpta Medica Foundation. 1968:456.
9. Hollingsworth PN, Cheah PS, Dawkins RL, *et al*. Observer variation in grading sacroiliac radiographs in HLA-B27 positive individuals. *J Rheumatol* 1983;10:247–54.
10. van Tubergen A, Heuft-Dorenbosch L, Schulpen G, *et al*. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? *Ann Rheum Dis* 2003;62:519–25.
11. van den Berg R, Lenczner G, Feydy A, *et al*. Agreement between clinical practice and trained central Reading in Reading of sacroiliac joints on plain pelvic radiographs: results from the DESIR cohort. *Arthritis Rheumatol* 2014;66:2403–11.
12. Bennett AN, McGonagle D, O'Connor P, *et al*. Severity of baseline magnetic resonance imaging–evident sacroiliitis and HLA–B27 status in early inflammatory back pain predict radiographically evident ankylosing spondylitis at eight years. *Arthritis Rheum* 2008;58:3413–18.
13. Feldtkeller E, Khan MA, van der Heijde D, *et al*. Age at disease onset and diagnosis delay in HLA-B27 negative vs. positive patients with ankylosing spondylitis. *Rheumatol Int* 2003;23:61–6.
14. Madsen KB, Jurik AG. Magnetic resonance imaging grading system for active and chronic spondylarthritis changes in the sacroiliac joint. *Arthritis Care Res* 2010;62:11–18.
15. Weber U, Lambert RGW, Pedersen SJ, *et al*. Assessment of structural lesions in sacroiliac joints enhances diagnostic utility of magnetic resonance imaging in early spondylarthritis. *Arthritis Care Res* 2010;62:1763–71.
16. Larbi A, Viala P, Molinari N, *et al*. Assessment of MRI abnormalities of the sacroiliac joints and their ability to predict axial spondyloarthritis: a retrospective pilot study on 110 patients. *Skeletal Radiol* 2014;43:351–8.
17. Dougados M, Etcheto A, Molto A, *et al*. Clinical presentation of patients suffering from recent onset chronic inflammatory back pain suggestive of spondyloarthritis: the DESIR cohort. *Jt Bone Spine Rev Rhum* 2015;82:345–51.
18. Calin A, Porta J, Fries JF, *et al*. Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977;237:2613–14.
19. Rudwaleit M, Metter A, Listing J, *et al*. Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. *Arthritis Rheum* 2006;54:569–78.
20. Weber U, Lambert RGW, Østergaard M, *et al*. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. *Arthritis Rheum* 2010;62:3048–58.
21. Baraliakos X, van der Heijde D, Braun J, *et al*. OMERACT magnetic resonance imaging initiative on structural and inflammatory lesions in ankylosing spondylitis—report of a special interest group at OMERACT 10 on sacroiliac joint and spine lesions. *J Rheumatol* 2011;38:2051–4.
22. Wick MC, Weiss RJ, Jaschke W, *et al*. Erosions are the most relevant magnetic resonance imaging features in quantification of sacroiliac joints in ankylosing spondylitis. *J Rheumatol* 2010;37:622–7.
23. Puhakka KB, Jurik AG, Egund N, *et al*. Imaging of sacroiliitis in early seronegative spondylarthropathy. Assessment of abnormalities by MR in comparison with radiography and CT. *Acta Radiol* 2003;44:218–29.
24. 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.
25. van den Berg R, Lenczner G, Thévenin F, *et al*. Classification of axial SpA based on positive imaging (radiographs and/or MRI of the sacroiliac joints) by local rheumatologists or radiologists versus central trained readers in the DESIR cohort. *Ann Rheum Dis* 2015;74:2016–21.
26. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37–46.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159.
28. Kundel HL, Polansky M. Measurement of observer agreement 1. *Radiology* 2003;228:303–8.
29. Weber U, Zubler V, Pedersen SJ, *et al*. Development and validation of a magnetic resonance imaging reference criterion for defining a positive sacroiliac joint magnetic resonance imaging finding in spondyloarthritis. *Arthritis Care Res* 2013;65:977–85.
30. de Hooge M, van den Berg R, Navarro-Compán V, *et al*. Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? *Ann Rheum Dis* 2016;75:1308–14.
31. Weber U, Pedersen SJ, Østergaard M, *et al*. Can erosions on MRI of the sacroiliac joints be reliably detected in patients with ankylosing spondylitis? A cross-sectional study. *Arthritis Res Ther* 2012;14:R124.
32. Poddubnyy D, Gaydukova I, Hermann KG, *et al*. Magnetic resonance imaging compared to conventional radiographs for detection of chronic structural changes in sacroiliac joints in axial spondyloarthritis. *J Rheumatol* 2013;40:1557–65.
33. Weber U, Pedersen SJ, Zubler V, *et al*. Fat infiltration on magnetic resonance imaging of the sacroiliac joints has limited diagnostic utility in nonradiographic axial spondyloarthritis. *J Rheumatol* 2014;41:75–83.
34. Krohn M, Braum LS, Sieper J, *et al*. Erosions and fatty lesions of sacroiliac joints in patients with axial spondyloarthritis: evaluation of different MRI techniques and two scoring methods. *J Rheumatol* 2014;41:473–80.
35. Madsen KB, Schiøttz-Christensen B, Jurik AG. Prognostic significance of magnetic resonance imaging changes of the sacroiliac joints in spondyloarthritis—a follow-up study. *J Rheumatol* 2010;37:1718–27.
36. Maksymowych WP, Wichuk S, Chiowchanwisawakit P, *et al*. Development and preliminary validation of the Spondyloarthritis Research Consortium of Canada magnetic resonance imaging sacroiliac joint structural score. *J Rheumatol* 2015;42:79–86.
37. Song IH, Hermann KG, Haibel H, *et al*. Inflammatory and fatty lesions in the spine and sacroiliac joints on whole-body MRI in early axial spondyloarthritis-3-Year data of the ESTHER trial. *Semin Arthritis Rheum* 2016;45:404–10.
38. Geijer M, Sihlbom H, Göthlin JH, *et al*. The role of CT in the diagnosis of sacro-iliitis. *Acta Radiol Stockh Swed 1987* 1998;39:265–8.
39. Guglielmi G, Scalzo G, Cascavilla A, *et al*. Imaging of the sacroiliac joint involvement in seronegative spondylarthropathies. *Clin Rheumatol* 2009;28:1007–19.
40. Heuft-Dorenbosch L, Landewé R, Weijers R, *et al*. Combining information obtained from magnetic resonance imaging and conventional radiographs to detect sacroiliitis in patients with recent onset inflammatory back pain. *Ann Rheum Dis* 2006;65:804–8.

RMD Open: first published as 10.1136/rmdopen-2016-000303 on 11 November 2016. Downloaded from http://rmdopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright.