Magnetic resonance imaging assessment of ASAS-defined active sacroiliitis in patients with inflammatory back pain and suspected axial spondyloarthritis: a study of reliability

L. Cereser[1], A. Zabotti[2], G. Zancan[1], L. Quartuccio[2], C. Cicciò[3], I. Giovannini[2], S. De Vita[2], C. Zuiani[1], R. Girometti[1]

[1]Institute of Radiology, Department of Medicine, [2]Rheumatology Clinic, University of Udine, University Hospital S. Maria della Misericordia, Udine, Italy; [3]Department of Diagnostic Imaging and Interventional Radiology, IRCCS Sacro Cuore Don Calabria Hospital, Negrar di Valpolicella, Verona, Italy.

Abstract

Objective

The main purpose was to investigate the intra- and inter-rater reliability of the Assessment of SpondyloArthritis international Society (ASAS) definition of positive MRI for active sacroiliitis (ASAS-positive MRI), in a sample of patients with inflammatory back pain (IBP) and suspected axial spondyloarthritis (axSpA), who underwent sacroiliac joints (SIJ) MRI. We also evaluated the intra- and inter-rater reliability for the detection of the recently ASAS-refined findings indicating inflammatory activity.

Methods

We retrospectively identified 105 consecutive patients with IBP and suspected axSpA who underwent SIJ MRI. Two radiologists in two distinct reading sessions assessed the prevalence of ASAS-positive MRI and of ASAS-defined signs of inflammatory activity. We determined the intra-rater and inter-rater reliability of the above-mentioned variables by means of prevalence-adjusted bias-adjusted kappa (PABAK) statistic, and verified whether there was any significant difference in providing the diagnosis of ASAS-positive MRI on an inter-rater basis (McNemar test).

Results

We observed substantial reliability in assessing a SIJ MRI as ASAS-positive both on intra-rater basis (PABAK ranging 0.70–0.77) and inter-rater basis (PABAK 0.71 for the first reading, and 0.64 for the second reading). No significant difference in the rate of diagnosis between raters was found ($p>0.99$ for both reading sets). Intra-rater and inter-rater reliability for inflammatory activity signs ranged from moderate to almost perfect.

Conclusion

The substantial intra- and inter-rater reliability in assessing the ASAS-positive MRI supports its use for classification purposes. The variable reliability of inflammatory activity signs suggests they are suboptimal as a complement to the current definition of ASAS-positive MRI.

Key words

magnetic resonance imaging; reliability, spondyloarthritis, sacroiliac joints

Lorenzo Cereser, MD
Alen Zabotti, MD
Giovanni Zancan, MD
Luca Quartuccio, MD, Prof.
Carmelo Cicciò, MD
Ivan Giovannini, MD
Salvatore De Vita, MD, Prof.
Chiara Zuiani, MD, Prof.
Rossano Girometti, MD, Prof.

Please address correspondence to:
Alen Zabotti,
Clinica di Reumatologia,
Università di Udine,
Ospedale S. Maria della Misericordia,
p.le S. Maria della Misericordia 15,
33100 Udine, Italy.
E-mail: alen.zabotti@asufc.sanita.fvg.it

Received on June 11, 2020; accepted in revised form on October 27, 2020.

Competing interests: none declared.

Introduction

Inflammatory back pain (IBP) is an insidious symptom with onset usually before 45 years of age, duration of at least 3 months, and improvement with exercise (1). The IBP is the peculiar clinical feature of axial spondyloarthritis (axSpA), with sensitivity between 70-80% and specificity varying according to the population (2). The term axSpA refers to a group of chronic immune-mediated inflammatory diseases affecting the axial skeleton. It primarily includes three clinical entities: non-radiographic axSpA (nr-axSpA), ankylosing spondylitis (AS), and psoriatic arthritis with axial involvement (ax-PsA). The diagnosis of axSpA derives from a comprehensive evaluation of clinical and laboratory data, in combination with imaging (3, 4). AS and nr-axSpA are probably two stages of the same disease involving the spine and the sacroiliac joints (SIJ). The former displays radiographic evidence of different stages of ankylosis of SIJ, while the latter, often considered as an early AS, displays sacroiliitis without radiographic damage. Lastly, ax-PsA is characterised by less severe and less symmetrical sacroiliitis, with clinical evidence of cutaneous psoriasis and lower frequency of HLA-B27 positivity (5).

Magnetic resonance imaging (MRI) proved to be the best imaging technique for the detection of SIJ signs of active inflammation, with accurate anatomic localisation, especially in the early phase of the disease (6). Currently, the detection of sacroiliitis on imaging (*i.e.* active inflammation on SIJs detected by MRI or definite radiographic sacroiliitis) is a major clinical trigger to start the biological disease-modifying anti-rheumatic drugs (b-DMARDs) in axSpA patients (7). The b-DMARDs proved to be the most effective treatment strategy to reach clinical remission and block disability quickly (7).

In this scenario, the Assessment of SpondyloArthritis international Society (ASAS) defined a standardised terminology for the qualitative description of active inflammatory and structural lesions detectable on MRI, according to an international consensus of experts (8, 9). A SIJ MRI score for quantifying inflammation (*i.e.* Spondyloarthritis Research Consortium of Canada – SPARCC MRI index) (10) has also been proposed. In the last ASAS update, the definition of "positive MRI for active sacroiliitis" was reaffirmed, while both the lesions indicating signs of inflammatory activity (*i.e.* bone marrow oedema, capsulitis, joint space enhancement, inflammation at the site of erosion, enthesitis, and joint space fluid) and the signs of structural change were partially revised and updated (11). In the same paper, Maksymowych *et al.* reported a preliminary validation of these definitions based on a selected cohort of patients (*i.e.* the ASAS classification cohort), while advocating further studies aimed to determine their usefulness in diagnosis, classification and prognosis, especially for clinical practice setting (11). To the best of our knowledge, no prior studies have validated the reproducibility of these MRI definitions yet.

The main purpose of the study was to investigate the intra- and inter-rater reliability of the ASAS definition of positive MRI for active sacroiliitis, in a sample of patients with IBP who underwent SIJ MRI for suspected axSpA. As secondary objectives, we evaluated the intra- and inter-rater reliability for: (i) the detection of MRI SIJ lesion ASAS definitions indicating signs of inflammatory activity; (ii) the attribution of the SPARCC MRI index for scoring the SIJ inflammation.

Materials and methods

Study population

Our Institutional Review Board approved the study. The need for written informed consent was waived due to the retrospective design.

By performing a computerised search, we identified all the consecutive adult patients with IBP and suspected axSpA who underwent SIJ MRI in our tertiary referral centre in the period July 2012 to January 2020. All patients were referred from the Rheumatology Clinic of the same centre. Concerning those patients who underwent multiple SIJ MRI, only the baseline examination was included in the analysis. On a total of 111 eligible patients, n=6 were excluded because of unavailable MRI examinations. There-

ASAS[1] definition of "positive MRI for active sacroiliitis"

MRI evidence of bone marrow edema, with:

- inflammation clearly present and located in the subchondral bone;
- appearance highly suggestive of spondyloarthritis;
- extension on ≥2 consecutive slices, or presence on a single slice if there is more than one inflammatory lesion.

ASAS definitions of SIJ[2] MRI inflammatory activity signs

Bone marrow edema: hyperintense signal on STIR, T2w Fat-Sat, or T1w Fat-Sat post-contrast images, relative to the sacral interforaminal bone marrow signal.

Inflammation at the site of erosion: hyperintense signal on STIR and/or T1w Fat-Sat post-contrast images at the site of erosion.

Capsulitis: hyperintense signal on STIR and/or T1w Fat-Sat post-contrast images at the perimeter of the joint (anterior or posterior and cranial or caudal on semi-axial and semi-coronal images, respectively).

Joint space fluid: hyperintense signal on STIR images in the joint space (equivalent to cerebrospinal fluid).

Joint space enhancement: hyperintense signal on T1w Fat-Sat post-contrast images in the joint space of the cartilaginous portion of the sacroiliac joint.

Enthesitis: hyperintense signal in bone marrow and/or soft tissue on STIR and/or T1w Fat-Sat post-contrast images at sites where ligaments and tendons attach to bone (but not including the inter-osseous ligaments of the sacroiliac joint).

[1]**ASAS**: Assessment of SpondyloArthritis international Society
[2]**SIJ**: sacroiliac joints

Modified from:
Maksymowych WP, et al. Ann Rheum Dis. 2019

Fig. 1. The definitions of "positive MRI for active sacroiliitis" and of MRI inflammatory activity signs according to ASAS.

fore, the final population included 105 patients (32 men and 73 women; median age, 46 years; interquartile range, 36–56 years). Thirty-five out of 105 (36.1%) had a confirmed clinical diagnosis of axSpA, including 11 nr-axSpA, 10 ax-PsA, and 14 AS.

SIJ MRI examinations

All SIJ MRI examinations were performed on a 1.5-Tesla equipment (Magnetom Avanto, Siemens Medical System, Erlangen, Germany), using a 32-channel surface coil. The standard MRI protocol included: semi-coronal (oriented on sacrum long axis) T1-weighted Turbo Spin-Echo (TSE) sequence (slice thickness, 3–4 mm; inter-slice gap, 10%; repetition time/echo time, 547/20 ms); semicoronal Short Tau Inversion Recovery (STIR) sequence (slice thickness, 3–4 mm; inter-slice gap, 10%; repetition time/echo time/time of inversion, 4780/94/150 ms); semi-axial STIR sequence (with the same parameters as above).

Imaging analysis and reliability exercise

The Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were followed for the preparation of the manuscript (12).

A study coordinator organised independent reading sessions to present the SIJ MRI images to two raters (R1 and R2), both radiologists with 10 years of experience in MRI. Each rater was blinded to the results of the other rater, as well as to patients' history, clinical data, and final diagnosis.

The study coordinator also provided readers a comprehensive imaging atlas, illustrating: (i) the ASAS definition of "positive MRI for active sacroiliitis", that is MRI evidence of bone marrow inflammation, with the following features: (a) bone marrow oedema (BME) on STIR sequences (or bone marrow contrast enhancement on a T1-weighted sequence) that is seen on at least two consecutive MRI slices or on a single slice if more than one inflammatory lesion is present; (b) inflammation clearly present and located in the subchondral bone; (c) MRI appearance highly suggestive of SpA; (ii) the new ASAS definitions of MRI findings indicating signs of inflammatory activity (*i.e.* subchondral BME, inflammation at the site of erosion, capsulitis, joint space enhancement, joint space fluid, and enthesitis) (11); (iii) how to calculate the SPARCC MRI index for scoring the SIJ inflammation (10). The definitions of "positive MRI for active sacroiliitis" and of MRI inflammatory activity signs according to ASAS are resumed in Figure 1. An example of SPARCC MRI index calculation is reported in Figure 2.

The two readers evaluated all the SIJ MRI examinations on a dedicated workstation (Olea Sphere, Olea Medical, La Ciotat, France) in two different reading sessions separated by a 4-week period, with examinations presented in different random orders, obtained using a freely available software on the Internet (https://www.randomizer.org). For each single reading, all the performed MRI sequences were disposable, giving the readers the possibility of simultaneous visualisation, in order to make proper anatomical and signal correlations.

Each reader, for each reading: (i) specified if the MRI examination was deemed positive or negative for active sacroiliitis according to ASAS definition; (ii) indicated the presence or absence of each of the aforementioned MRI lesions indicating signs of inflammatory activity; (iii) calculated the SPARCC MRI index for scoring the SIJ inflammation (score range, 0–72) (10, 11). Of note, analysis did not include "joint space enhancement" since no patients underwent contrast medium administration.

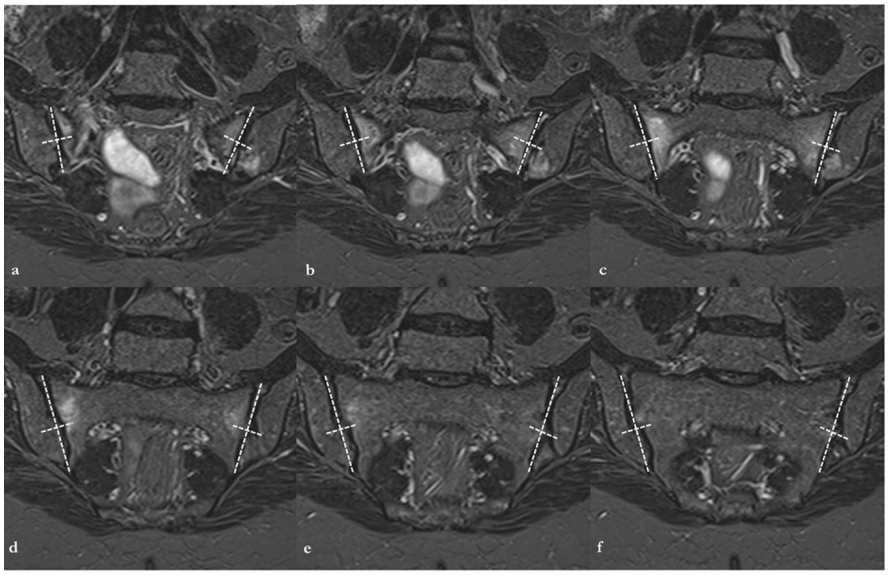

Fig. 2. SPARCC MRI index with total score of 31, calculated in a 46-year-old woman with ankylosing spondylitis. On the selected six semi-coronal STIR images **A-F**: 21 points were granted for BME (example in **b**, 2 points for segments of the right SIJ and 3 points for segments of the left SIJ); 5 points were granted for BME depth >1 cm measured perpendicular to the articular surface (example in **c**, 1 point in the right sacral wing); 5 points were granted for BME intensity (example in **c**, 1 point in the right sacral wing).
SPARCC: Spondyloarthritis Research Consortium of Canada; STIR: short tau inversion recovery; BME: bone marrow oedema; SIJ: sacroiliac joint.

Table I. Prevalence of ASAS-positive MRI examinations and of inflammatory activity signs.

	1st Reading		2nd Reading	
	R1 n (%)	**R2** n (%)	**R1** n (%)	**R2** n (%)
ASAS-positive MRI	25 (23.8)	26 (24.8)	29 (27.6)	28 (26.7)
Subchondral bone marrow oedema	49 (46.7)	42 (40.0)	46 (43.8)	39 (37.1)
Inflammation at the site of erosion	19 (18.1)	15 (14.3)	18 (17.1)	24 (22.9)
Capsulitis	3 (2.9)	10 (9.5)	5 (4.8)	7 (6.7)
Joint space fluid	20 (19.0)	10 (9.5)	19 (18.1)	9 (8.6)
Enthesitis	19 (18.1)	10 (9.5)	18 (17.1)	8 (7.6)

R1: rater 1; R2: rater 2.

Statistical analysis

We used Percent Agreement (PA) and Cohen's Kappa (k) with 95% confidence intervals (CI) to determine the intra-rater and inter-rater reliability for the nominal categorical variables (*i.e.* the ASAS definition of positive MRI for active sacroiliitis, and each single ASAS definition indicating signs of inflammatory activity). When paradox k was observed (*i.e.* acceptable PA, and unacceptable k) and both Prevalence Index and Bias Index were different from zero, the imbalance was corrected by using the prevalence-adjusted bias-adjusted kappa (PABAK) statistic (13-15). Interpretation of k and PABAK coefficient was as follow: <0.00, poor; 0.00–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; 0.81–1.00, almost perfect (16). We used the McNemar test to verify whether there was any significant difference in providing the diagnosis of positive MRI on an inter-rater basis. Intra-rater and inter-rater reliability of the SPARCC MRI index (*i.e.* continue variable) was assessed with the intraclass correlation coefficients (ICC). As ICC models we used the "two-way mixed effects, absolute agreement, single rater/measurement" for the intra-rater reliability, and the "two-way random effects, absolute agreement, single rater/measurement" for the inter-rater reliability (17). Based on the 95% CI of the ICC estimates, values <0.50, 0.50–0.75, 0.75–0.90, and >0.90 were considered indicative of poor, moderate, good, and excellent reliability, respectively (18). We used the Wilcoxon signed-rank test to verify whether there was a significant difference between the SPARCC MRI index, as calculated by R1 *versus* R2.

The reference α value was 0.05. All statistical analyses were performed using commercially available software (MedCalc Software bvba, v. 18.11.6, Ostend, Belgium).

Results

MRI findings

All the 105 SIJ MRI examinations included both T1-weighted TSE and STIR sequences in the semi-coronal plane, with complete visualisation of both SIJ in all cases.

The per-reader prevalence of ASAS-positive MRI diagnoses and inflammatory activity signs is reported in Table I. Overall, the prevalence of ASAS-positive MRI diagnoses was comparable on intra- and inter-rater basis, ranging 24-25% in the first readings set, and 27–28% in the second readings set, respectively. By averaging the results on an intra- and inter-rater basis, we observed a 26% mean prevalence of ASAS-positive MRI diagnoses (example case in Fig. 3), 42% of subchondral BME, 18% of inflammation at the site of erosion, 6% of capsulitis, 14% of joint space fluid, and 13% of enthesitis.

Reliability of qualitative imaging findings

Supplementary Table I and Supplementary Table S2 show the results for the PA and Cohen's kappa, from which we derived PABAK values presented in Table II (intra-rater reliability) and Table III (inter-rater reliability). Supplementary Tables S3 and S4 illustrate the distribution of ASAS definition of "positive MRI" and of inflammatory activity signs on a per-reader basis in the first reading and in the second reading, respectively.

Concerning intra-rater evaluation (Table II), we observed substantial reliabil-

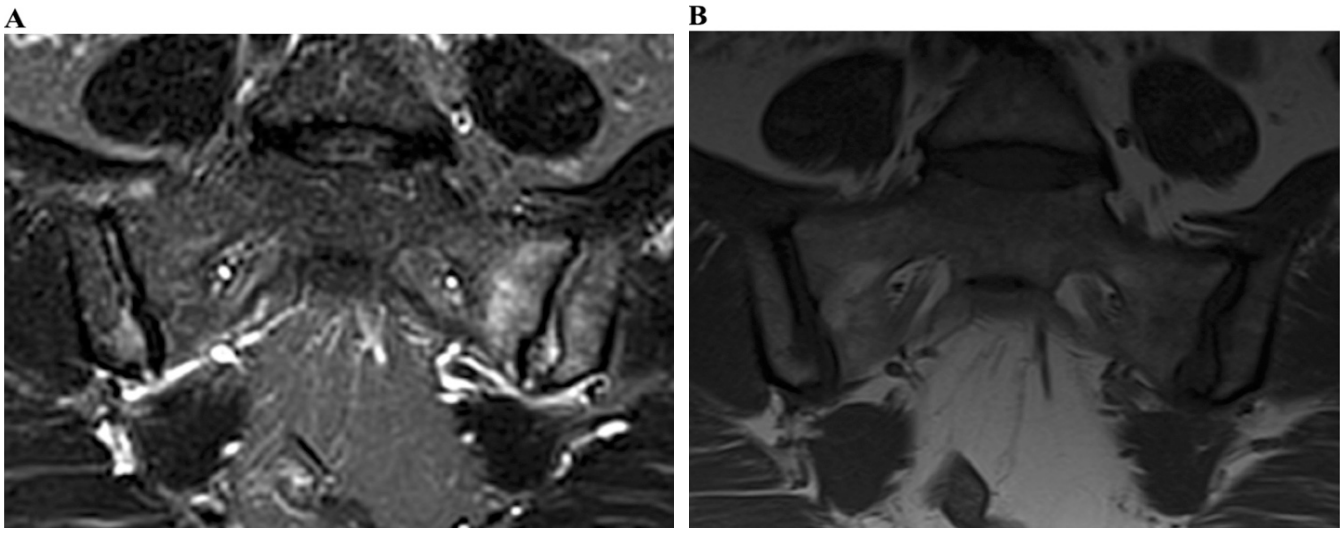

Fig. 3. ASAS positive MRI for active sacroiliitis in a 58-year-old woman with non-radiographic axSpA. Semi-coronal STIR image (**A**) and corresponding semi-coronal TSE T1-weighted image (**B**) show subchondral bone marrow oedema in both sides of the left SIJ and in the right inferior iliac quadrant. ASAS: Assessment of SpondyloArthritis international Society; MRI: magnetic resonance imaging; axSpA: axial spondyloarthritis; STIR: short tau inversion Recovery; TSE: turbo spin-echo; SIJ: sacroiliac joint.

ity in assessing a case as ASAS-positive (PABAK ranging 0.70–0.77). Intra-rater reliability for inflammatory activity signs ranged from moderate to almost perfect, with the higher reliability in the case of capsulitis (PABAK 0.89 and 0.90 for R1 and R2, respectively) and joint space fluid (PABAK 0.79 and 0.98 for R1 and R2, respectively).

On an inter-rater basis, we observed substantial reliability in assessing a case as ASAS-positive both in the first reading (PABAK 0.71) and in the second reading (PABAK 0.64), with no significant difference in the rate of diagnosis (p>0.99 for both reading sets). Inter-rater reliability for inflammatory activity signs ranged from moderate to almost perfect both in the first and in the second reading. Higher reliability was found in the case of capsulitis, regardless of the reading set.

Quantitative imaging findings and related reliability

SPARCC MRI index values (Table IV) showed no significant difference between R1 and R2 in the first reading set (p=0.23) and second reading set (p=0.16).

The reliability for the index was excellent on an intra-rater basis (ICC of 0.96 [95%CI 0.94–0.97] for R1, and ICC of 0.97 [95%CI 0.95–0.98] for R2), and good-to-excellent on an inter-rater basis (ICC of 0.89 [95%CI 0.84–0.92] for the first reading set, and 0.81 [95%CI 0.73–0.87] for the second reading set).

Discussion

In our cohort of patients with IBP, the reliability for the assessment of a SIJ MRI case as ASAS-positive among two experienced radiologists was substantial, both on an intra-rater and on an inter-rater basis. Secondarily, the reliability for the detection of the MRI-detected inflammatory activity signs was moderate to almost perfect.

We found a mean prevalence of ASAS-positive MRI of 26%, similar to Maksymowych *et al*. on the ASAS classification cohort patients (31%) (11) and also in line with previous studies, in which reported prevalences ranged from 21% (19) to 41% (20). Subchondral BME was the most frequently reported inflammatory activity sign, with a mean prevalence (of 42%) that is equivalent to other studies (*e.g*. Maksymowych *et al*., 40%; Jans *et al*., 42%) (11, 21). For both raters, the ASAS definition of positive MRI was attributed to a smaller number of patients compared to the total cases with subchondral BME, indicating that raters discerned between BME indicative of active sacroiliitis and BME not complying to the ASAS definition of positive MRI. The intra-rater and inter-rater reliability obtained for the ASAS definition of positive MRI was substantial, in accordance with previous studies (11, 22), thus confirming its primary role in the clinical practice for diagnosis and classification purposes. Of note, recently BME in the SIJ according to the ASAS definition of positive MRI was detected in 17% of healthy volunteers (23). Patients' age and Body Mass Index were associated with the MRI detection of BME, highlighting a possible influence of osteoarthritis, physical activity, and mechanical factors as causes of BME and ASAS-positive MRI (24, 25). This high frequency of active inflammatory SIJ MRI findings suggestive for axSpA suggests a need for an update for the diagnosis and classification criteria of axSpA, in order to improve the specificity of the imaging criterion.

As expected from the low frequency of axSpA in the tested sample, all other MRI inflammatory activity signs were rarer, although they showed nearly double frequency than Maksymowych *et al*. (*i.e*. inflammation at the site of erosion, 18% *vs*. 7%; capsulitis, 6% *vs*. 3%; joint space fluid, 14% *vs*. 7%; enthesitis 13% *vs*. 5%). This discrepancy may be related to intrinsic differences in the cohorts of patients and/or in the selection of raters. Since the ASAS definitions of most of the above-mentioned MRI signs are new or recently renewed, further studies are needed to estimate the real prevalences both in patients and in healthy controls. Capsu-

Table II. Intra-rater reliability of the ASAS definition of "positive MRI" and of inflammatory activity signs.

	R1 PABAK (95%C.I.)	R2 PABAK (95%C.I.)
ASAS-positive MRI	0.77 (0.61-0.93)	0.70 (0.52-0.88)
Subchondral bone marrow oedema	0.75 (0.63-0.88)	0.71 (0.57-0.86)
Inflammation at the site of erosion	0.49 (0.20-0.77)	0.68 (0.45-0.91)
Capsulitis	0.89 (0.28-1.00)	0.90 (0.63-1.00)
Joint space fluid	0.79 (0.59-0.99)	0.98 (0.87-1.00)
Enthesitis	0.64 (0.38-0.89)	0.92 (0.69-1.00)

R1: rater 1; R2: rater 2.

Table III. Inter-rater reliability of the ASAS definition of "positive MRI" and of inflammatory activity signs.

	1st Reading PABAK (95%C.I.)	2nd Reading PABAK (95%C.I.)
ASAS-positive MRI	0.71 (0.53-0.90)	0.64 (0.45-0.82)
Subchondral bone marrow oedema	0.52 (0.36-0.69)	0.56 (0.40-0.73)
Inflammation at the site of erosion	0.54 (0.25-0.84)	0.54 (0.29-0.79)
Capsulitis	0.83 (0.38-1.00)	0.89 (0.47-1.00)
Joint space fluid	0.70 (0.42-0.97)	0.81 (0.57-1.00)
Enthesitis	0.56 (0.23-0.89)	0.66 (0.33-0.98)

Table IV. SPARCC MRI index for scoring sacroiliac joint inflammation.

	1st Reading		2nd Reading	
	R1	R2	R1	R2
Mean value	3.8	4.6	3.4	4.6
Standard deviation	8.4	10.2	7.1	10.5

R1: rater 1; R2: rater 2.

litis was the less frequent sign; the real prevalence may be underestimated, due to the lack of contrast-enhanced MRI examinations, which may have revealed capsulitis in a higher number of cases in the form of soft tissue enhancement at the perimeter of the SIJ. However, our institutional policy not to routinely use contrast medium derives from European Society of Skeletal Radiology (ESSR) and European League Against Rheumatism (EULAR) recommendations, stating that STIR sequences are generally sufficient to detect SIJ inflammation (26, 27).

We believe that the difference in reliability for the detection of subchondral BME between the intra-rater (substantial) and the inter-rater evaluation (moderate) may be due to individual reporting style differences, presumably related to different attribution of ambiguous, small signal hyperintensity areas on STIR images (small amount of BME *vs.* artifact). The moderate inter-rater reliability for subchondral BME we obtained is in line with previous results (28). Nevertheless, as reported above, moderate reliability for subchondral BME translated into substantial reliability for the ASAS definition of "positive MRI", which is the primary task for radiologists evaluating SIJ MRI. Reliability for inflammation at the site of erosion and enthesitis was only moderate. This may be due to recent introduction or revision of those signs (11), which in turn might have increased the risk of errors in evaluation. In addition, the inflammation at the site of erosion is difficult to assess, since its identification requires a simultaneous and meticulous visual analysis of both T1-weighted and STIR MR images. According to the ASAS definition, enthesitis is defined as the presence of bright signal on STIR images in the bone marrow of the iliac bone, in a site that is posterior to the SIJ (11). A previous study on pelvic enthesitis at MRI (29) assessed that soft tissue inflammation was far more common than BME in enthesitis, thus presumably leading to a lower detection rate and reliability. Although capsulitis and joint space fluid were infrequent findings in our sample, reliability for their detection ranged from substantial to almost perfect. Nevertheless, the clinical usefulness of these signs is limited to a supportive role in diagnosing active sacroiliitis when typical subchondral BME is present.

The SPARCC MRI index aims to quantify the subchondral BME in the synovial portion of the SIJ, given the number of slices and quadrants in which it is visible, with additional points depending on its intensity and depth (10). It was developed for classification purposes and patient allocation in the field of research (30, 31), with little role in every-day clinical practice (32). The excellent reliability we found on intra-rater and inter-rater bases allows us to consider the SPARCC MRI index a repeatable and reproducible tool in scoring the SIJ inflammation, thus confirming its role in research studies concerning objective quantification of BME. Our results are in line with previous studies, in which high reliability of the SPARCC MRI index was found in the setting of both adult (10, 28, 33) and paediatric population with suspected or confirmed spondyloarthritis (34, 35).

Our study had some limitations. First, we found a low prevalence of axSpA, thus presumably conditioning a low prevalence of all the MRI inflammatory signs other than BME. However, this is in line with a previous study, demonstrating axSpA in 35% of patients with IBP (36). Therefore, further studies assessing the reliability of MRI inflammatory signs in cohorts of patients with axSpA are needed. Second, since we did not involve raters with different experience in MRI, we could not assess the influence of the experience on reliability. Nevertheless, we involved experienced raters, suggesting that our results can be generalisable to tertiary referral centres. Third, we did not evaluate whether the reliability we assessed translated into better diagnostic accura-

cy or more effective patients' management, though it is reasonable to assume that substantial reliability has the potential to affect clinical decisions (37).

In conclusion, we found substantial intra- and inter-rater reliability in assessing positive SIJ MRI with ASAS criteria, thus supporting their use for classification purposes. The reliability of the inflammatory activity signs was very variable, ranging from moderate to almost perfect, suggesting they are suboptimal as a mean to complement the current definition of ASAS-positive MRI in clinical practice.

References

1. CALIN A, PORTA J, FRIES JF, SCHURMAN DJ: Clinical history as a screening test for ankylosing spondylitis. *JAMA* 1977; 237: 2613-4.
2. SIEPER J, VAN DER HEIJDE D, LANDEWÉ R *et al*: New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). *Ann Rheum Dis* 2009; 68: 784-8.
3. LEDFORD C: Spine conditions: mechanical and inflammatory low back pain. *FP Essent* 2017. 461: 15-20.
4. NADON V, MOLTÓ A, ETCHETO A *et al*: Clinical peripheral enthesitis in the DESIR prospective longitudinal axial spondyloarthritis cohort. *Clin Exp Rheumatol*. 2019; 37: 561-5.
5. HELLIWELL PS: Axial disease in psoriatic arthritis. *Rheumatology* (Oxford) 2020; 59: 1193-5.
6. OOSTVEEN J, PREVO R, DEN BOER J, VAN DE LAAR M: Early detection of sacroiliitis on magnetic resonance imaging and subsequent development of sacroiliitis on plain radiography. A prospective, longitudinal study. *J Rheumatol* 1999; 26: 1953-8.
7. VAN DER HEIJDE D, RAMIRO S, LANDEWÉ R *et al*: 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis* 2017; 76: 978-91.
8. RUDWALEIT M, JURIK AG, HERMANN KG *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.
9. LAMBERT RG, BAKKER PA, VAN DER HEIJDE D *et al:* Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. *Ann Rheum Dis* 2016; 75: 1958-63.
10. MAKSYMOWYCH WP, INMAN RD, SALONEN D *et al*: Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. *Arthritis Rheum* 2005; 53: 502-9.
11. MAKSYMOWYCH WP, LAMBERT RG, ØSTERGAARD M *et al*: MRI lesions in the sacroiliac joints of patients with spondyloarthritis: an update of definitions and validation by the ASAS MRI working group. *Ann Rheum Dis* 2019; 78: 1550-8.
12. KOTTNER J, AUDIGÉ L, BRORSON S *et al*: Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011; 64: 96-106.
13. CICCHETTI DV, FEINSTEIN AR: High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990; 43: 551-8.
14. BYRT T, BISHOP J, CARLIN JB: Bias, prevalence and kappa. *J Clin Epidemiol* 1993; 46: 423-9.
15. NURJANNAH I, SIWI SM: Guidelines for analysis on measuring interrater reliability of nursing outcome classification. *Int J Res Med Sci* 2017; 5: 1169-75.
16. LANDIS JR, KOCH GG: The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74.
17. HALLGREN KA: Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol* 2012; 8: 23-34.
18. KOO TK, LI MY: A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 2016; 15: 155-63.
19. TARANTINO A, JABLONSKA JP, D'APRILE P: All that glitters is not gold: sacroiliitis. *Reumatologia* 2018; 56: 289-93.
20. SUDOŁ-SZOPIŃSKA I, KWIATKOWSKA B, WŁODKOWSKA-KORYTKOWSKA M, MATUSZEWSKA G, GROCHOWSKA E: Diagnostics of sacroiliitis according to ASAS criteria: a comparative evaluation of conventional radiographs and MRI in patients with a clinical suspicion of spondyloarthropathy. preliminary results. *Pol J Radiol* 2015; 80: 266-76.
21. JANS L, COEMAN L, VAN PRAET L *et al*: How sensitive and specific are MRI features of sacroiliitis for diagnosis of spondyloarthritis in patients with inflammatory back pain? *JBR-BTR* 2014; 97: 202-5.
22. 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.
23. BARALIAKOS X, RICHTER A, FELDMANN D *et al*: Frequency of MRI changes suggestive of axial spondyloarthritis in the axial skeleton in a large population-based cohort of individuals aged <45 years. *Ann Rheum Dis* 2020; 79: 186-92.
24. BRAUN J, BARALIAKOS X, BUEHRING B, KILTZ U, FRUTH M: Imaging of axial spondyloarthritis. New aspects and differential diagnoses. *Clin Exp Rheumatol*. 2018; 36 (Suppl. 114): S35-42.
25. ESATOGLU SN, HATEMI G: Bone marrow oedema on sacroiliac/spine MRI: is it really a sign of objective inflammation warranting treatment? *Clin Exp Rheumatol*. 2019; 37 (Suppl. 122): S10.
26. SCHUELLER-WEIDEKAMM C, MASCARENHAS VV, SUDOL-SZOPINSKA I *et al*: Imaging and interpretation of axial spondylarthritis: the radiologist's perspective--consensus of the Arthritis Subcommittee of the ESSR. *Semin Musculoskelet Radiol* 2014; 18: 265-79.
27. 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.
28. DE WINTER J, DE HOOGE M, VAN DE SANDE M: Magnetic resonance imaging of the sacroiliac joints indicating sacroiliitis according to the assessment of spondyloarthritis international society definition in healthy individuals, runners, and women with postpartum back pain. *Arthritis Rheumatol* 2018; 70: 1042-48.
29. HERREGODS N, DEHOORNE J, PATTYN E *et al*: Diagnositic value of pelvic enthesitis on MRI of the sacroiliac joints in enthesitis related arthritis. *Pediatr Rheumatol Online J* 2015; 13: 46.
30. SIEPER J, VAN DER HEIJDE D, DOUGADOS M *et al*: A randomized, double-blind, placebo-controlled, sixteen-week study of subcutaneous golimumab in patients with active nonradiographic axial spondyloarthritis. *Arthritis Rheumatol* 2015; 67: 2702-12.
31. MAKSYMOWYCH WP: Imaging in axial spondyloarthritis: evaluation of inflammatory and structural changes. *Rheum Dis Clin North Am* 2016; 42: 645-62.
32. TSOI C, GRIFFITH JF, LEE RKL, WONG PCH, TAM LS: Imaging of sacroiliitis: Current status, limitations and pitfalls. *Quant Imaging Med Surg* 2019; 9: 318-35.
33. MAKSYMOWYCH WP, LAMBERT RG, BROWN LS, PANGAN AL: Defining the Minimally Important Change for the SpondyloArthritis Research Consortium of Canada Spine and Sacroiliac Joint Magnetic Resonance Imaging Indices for Ankylosing Spondylitis. *J Rheumatol* 2012; 39: 1666-74.
34. WEISS PF, COLBERT RA: Juvenile spondyloarthritis: a distinct form of juvenile arthritis. *Pediatr Clin North Am* 2018; 65: 675-90.
35. PANWAR J, TSE SML, LIM L *et al*: Spondyloarthritis Research Consortium of Canada scoring system for sacroiliitis in juvenile spondyloarthritis/enthesitis-related arthritis: a reliability, validity, and responsiveness Study. *J Rheumatol* 2019. 46: 636-44.
36. BRAUN A, SARACBASI E, GRIFKA J, SCHNITKER J, BRAUN J: Identifying patients with axial spondyloarthritis in primary care: how useful are items indicative of inflammatory back pain? *Ann Rheum Dis* 2011; 70: 1782-87.
37. KRAEMER HC: The reliability of clinical diagnoses: state of the art. *Annu Rev Clin Psychol* 2014; 10: 111-30.