

# Development and Preliminary Validation of the Spondyloarthritis Research Consortium of Canada Magnetic Resonance Imaging Sacroiliac Joint Structural Score

Walter P. Maksymowych, Stephanie Wichuk, Praveena Chiowchanwisawakit, Robert G. Lambert, and Susanne J. Pedersen

**ABSTRACT. Objective.** There is an unmet need for reliable assessment of structural progression in the sacroiliac joints (SIJ) of patients with spondyloarthritis (SpA), but radiography is unreliable and lacks responsiveness. We aimed to develop and validate a new scoring method for structural lesions based on magnetic resonance imaging (MRI), the Spondyloarthritis Research Consortium of Canada (SPARCC) SIJ Structural Score (SSS).

**Methods.** The SSS method for assessment of structural lesions is based on T1-weighted spin echo MRI, validated lesion definitions, slice selection according to well-defined anatomical principles, and dichotomous scoring (lesion present/absent) of 5 consecutive slices through the cartilaginous portion of the joint. Scoring ranges are fat metaplasia (0–40), erosion (0–40), backfill (0–20), and ankylosis (0–20). We progressively conducted 3 validation exercises with 2–4 readers on baseline, and either 2-year (exercises 1 and 2) or 1-year (exercise 3) scans from 147 patients with SpA assessed blinded to timepoint. Interobserver reliability was assessed by intraclass correlation coefficient (ICC) and smallest detectable change (SDC).

**Results.** Interobserver reliability for status score was good to excellent for ankylosis (ICC 0.79–0.98), consistently good for fat metaplasia (ICC 0.71–0.78), moderate to good for erosion (ICC 0.58–0.62), and fair to good for backfill (ICC 0.35–0.66). Reliability for change scores was moderate to good for all structural lesions despite the relatively small changes in scores, and was highest for fat metaplasia when both ICC and SDC values were compared.

**Conclusion.** The new SPARCC MRI SSS method can detect structural changes in the SIJ with acceptable reliability over a 1–2-year timeframe, and should be further validated in patients with SpA. (First Release Oct 15 2014; J Rheumatol 2015;42:79–86; doi:10.3899/jrheum.140519)

## Key Indexing Terms:

MRI                      VALIDATION                      SACROILIAC JOINT                      SPONDYLOARTHTRITIS

---

From the Department of Medicine; Department of Radiology and Diagnostic Imaging, Spondyloarthritis Research Consortium of Canada (SPARCC) Center, University of Alberta, Edmonton, Alberta, Canada; the Department of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand; the Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, University of Copenhagen, Denmark. Supported by a research fellowship grant from the SPARCC and a grant from the Danish Council for Independent Research in the Medical Sciences to Dr. Pedersen. Dr. Maksymowych is a Medical Scientist of Alberta Innovates Health Solutions.

W.P. Maksymowych, FRCP(C), Professor; S. Wichuk, BSc, Research Associate, Department of Medicine; R.G. Lambert, FRCP(C), Professor, Department of Radiology and Diagnostic Imaging, Spondyloarthritis Research Consortium of Canada Center, University of Alberta; P. Chiowchanwisawakit, MD, Department of Medicine, Siriraj Hospital, Mahidol University; S.J. Pedersen, MD, Rheumatologist, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, University of Copenhagen.

Address correspondence to Dr. W.P. Maksymowych, Professor of Medicine, University of Alberta, 562 Heritage Medical Research Building, Edmonton, Alberta T6G 2S2, Canada.

E-mail: walter.maksymowych@ualberta.ca

Accepted for publication September 2, 2014.

Radiography of the sacroiliac joints (SIJ) continues to be an important diagnostic tool in patients with spondyloarthritis (SpA). However, it is unreliable for detecting change and lacks responsiveness so it is not used for assessment of potential disease-modifying agents. There is, therefore, an unmet need for imaging tools to assess the potential disease-modifying effects of therapeutic agents early in SpA when disease is still confined to the SIJ<sup>1,2,3</sup>.

Magnetic resonance imaging (MRI) represents a substantial advance in the field because of its ability to visualize inflammation in soft tissue as well as the subchondral bone, visible as bone marrow edema (BME)<sup>4</sup>. This is evident on fat-suppressed sequences such as short-tau inversion recovery (STIR). Recent MRI data also show that resolution of BME in subchondral bone marrow may be associated with the development of fat metaplasia on the T1-weighted spin echo (T1WSE) sequence<sup>1,5,6</sup>. Fat metaplasia is not observed on radiography and the

histopathology of this lesion is unknown, but it is frequently observed in the subchondral marrow of SIJ and at spinal locations that are also typical for inflammation, i.e., vertebral corners, adjacent to vertebral endplates, and facet and costovertebral joints<sup>7</sup>. Moreover, fat metaplasia at vertebral corners has been shown to predict the subsequent development of new syndesmophytes in multivariate analysis<sup>1</sup>. We have previously hypothesized that the resolution of inflammation in SIJ erosions is followed by the development of a new tissue, which on T1WSE MRI has the same signal intensity as fat metaplasia<sup>8</sup>. We have called this type of lesion “backfill” because of its appearance in the cavity of the erosion, whereas the term fat metaplasia is used when this lesion is located in the subchondral bone marrow.

There has been limited validation of MRI-based scores for structural lesions in the SIJ, primarily based on cross-sectional data and recording fat metaplasia<sup>6,8,9,10,11,12,13,14</sup>, but it is unclear whether change in structural progression can be reliably detected. The Outcome Measures in Rheumatology (OMERACT) filter is a well-accepted framework for the validation of new scoring instruments based on assessment of feasibility, truth, and discrimination<sup>15</sup>. In our report, we describe the development and preliminary validation of a new scoring method for structural lesions in the SIJ according to the OMERACT filter, the Spondyloarthritis Research Consortium of Canada (SPARCC) Sacroiliac Joint Structural Score (SSS). This method assesses a broader spectrum of structural lesions in the SIJ that includes erosion, fat metaplasia, backfill, and ankylosis. Preliminary validation of this scoring method was aimed at demonstrating reliability of detection of structural lesions in the SIJ, at detection of change in structural progression in the SIJ, and whether change could be reliably detected before 2 years, the time frame required before radiographic progression in the spine can be reliably detected<sup>16</sup>.

## MATERIALS AND METHODS

**Patients.** We assessed MRI scans from patients meeting the modified New York criteria for ankylosing spondylitis recruited to the SPARCC prospective observational cohort<sup>17</sup>. Patients are evaluated systematically every 6 months according to a standardized protocol that includes clinical and laboratory variables. Imaging variables included SPARCC MRI SIJ inflammation scores<sup>18</sup> that are recorded for each patient in the cohort by readers unconnected with our study at baseline, at 3–6 months for patients starting antitumor necrosis factor- $\alpha$  (anti-TNF- $\alpha$ ), and annually for all patients as described<sup>17</sup>. A total of 147 patients had available baseline and 2-year followup scans, and were included in our study. Our study was performed in accordance with the Helsinki Declaration. A written informed consent is obtained from all study participants before inclusion into the observational cohort.

**MRI protocol.** The SPARCC SSS method was developed based on semicoronal T1WSE sequences of the SIJ. The scan variables were as follows: 15–19 slices, 4-mm slice thickness, 0.4-mm interslice gap, field of view 280–300 mm, repetition time 423–450 ms, echo time 12–13 ms, and matrix 512  $\times$  256 pixels. Although the original scans from all patients included STIR sequences, the STIR sequences were deleted from the set of scans included in this validation process to ensure blinding of the readers

with respect to timepoint and change in the distribution of active inflammation. In particular, the substantial reduction in BME that is typically observed in patients receiving anti-TNF- $\alpha$  therapy may unblind readers to time sequence.

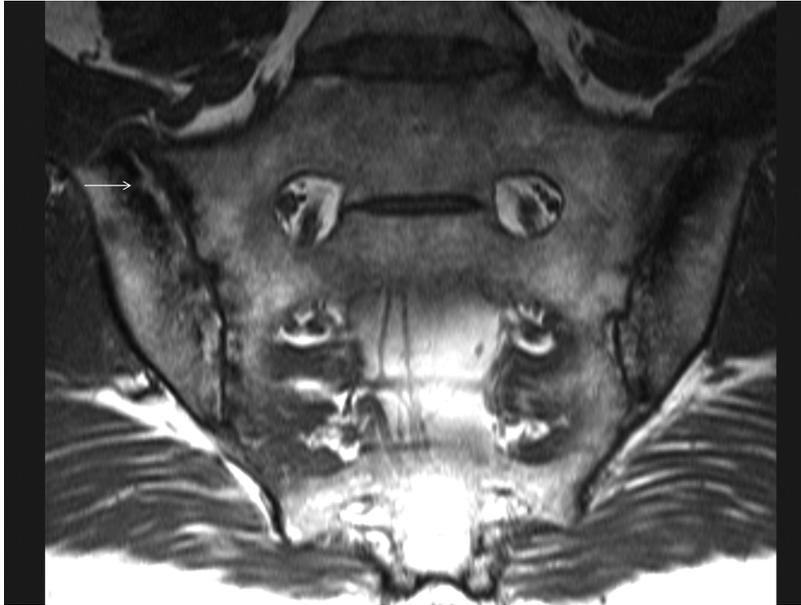
**Development of the SPARCC SSS method. Lesion definitions.** We adopted standardized definitions of structural lesions of the SIJ on MRI, which were developed by the Canada-Denmark MRI Working Group<sup>19</sup> and then extended in a subsequent report to include backfill<sup>8</sup>. For fat metaplasia to be scored in the SSS method, the lesion has to demonstrate homogeneous signal across the lesion that must extend more than 1 cm in depth from the joint surface because our previous work has shown that this enhances reliability (data not shown). Backfill is defined on T1WSE sequences as the complete loss of the iliac or sacral cortical bone at its anticipated location and increased signal that is clearly demarcated from adjacent normal marrow by irregular dark signal reflecting sclerosis at the border of the eroded bone (Figure 1).

We developed a training module that comprised a detailed description of the SSS scoring method with examples, a Digital Imaging and Communication in Medicine (DICOM)-based reference image set of 45 cases with baseline and 2-year scans, and consensus reader scores for these DICOM images. This comprehensive module is aimed at facilitating calibration of non-expert readers and is available at [www.carearthritis.com](http://www.carearthritis.com). Bone sclerosis and abnormalities of the synovial cavity are not addressed in the SSS method because of poor reproducibility in previous reading exercises (data not shown).

**Scoring methodology.** Evaluation of structural lesions in the SIJ proceeds sequentially in the following steps: (1) the transitional slice is identified by scrolling from anterior to posterior through the SIJ and viewing DICOM images depicting semicoronal slices through the joint. The transitional slice is defined as the first slice in the cartilaginous portion that has a visible portion of the ligamentous joint when viewed from anterior to posterior (Figure 2); (2) all timepoints are anatomically matched according to the transitional semicoronal SIJ slice. The link function on the DICOM viewing software allows simultaneous scrolling of anatomically matched images from the transitional slice anteriorly, thereby facilitating detection of change in lesions between different timepoints; (3) five consecutive semicoronal slices are assessed starting from the transitional slice and scrolling anteriorly. This number of slices was chosen because our preliminary work showed that the SIJ cavity together with adjacent bone marrow are still clearly visible at the most anterior slice in virtually all patients (data not shown). Scoring slices posterior to the transitional slice is not appropriate because those slices are predominantly or entirely composed of the ligamentous portion of the joint, which often has irregular bone cortices resembling erosions. Consequently, additional anterior or posterior slices do not provide useful additional information, depict areas that are frequently the hardest to interpret even in normal subjects, and including them in the exercise reduced reader reliability; and (4) the presence/absence of lesions is scored in SIJ quadrants (fat, erosion) or halves (backfill, ankylosis) using a direct online data entry system based on a schematic of the SIJ.

Scoring ranges are fat metaplasia (0–40), erosion (0–40), backfill (0–20), and ankylosis (0–20). The scoring for erosion and fat metaplasia is based on the assessment of SIJ quadrants because those abnormalities can occur on either the iliac and/or sacral sides. Scoring of those lesions is therefore based on all 8 SIJ quadrants. Ankylosis extends from the iliac to the sacral side while backfill fills the joint cavity. Scoring of these lesions is therefore based on SIJ halves. An example of the approach to scoring is provided in Figure 3.

**Reading exercises.** We conducted 3 formal reading exercises aimed at the assessment of feasibility and reliability. Readings were conducted blinded to patient demographics, time sequence, and treatment according to the following steps: (1) feasibility was first assessed in a pilot exercise of 20 cases randomly selected from the cohort with baseline and 2-year scans scored by 4 readers blinded to timepoint. The average time required to read



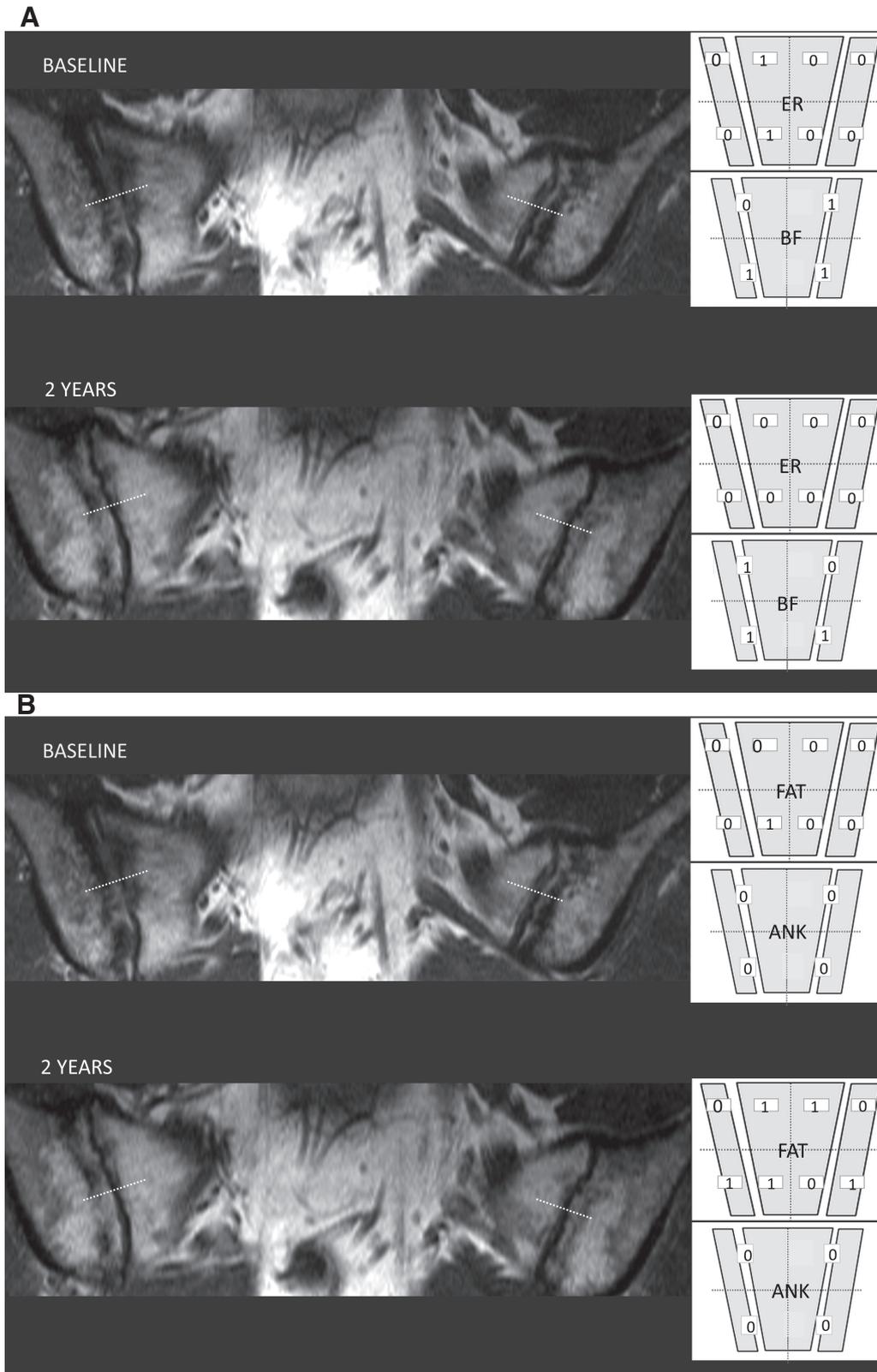
*Figure 1.* T1WSE MRI scan illustrating an example of backfill. There is complete loss of the right iliac cortical bone at its anticipated location and increased signal (white arrow) that is clearly demarcated from adjacent normal marrow by irregular dark signal, reflecting sclerosis at the border of the eroded bone. T1WSE: T1-weighted spin echo; MRI: magnetic resonance imaging.



*Figure 2.* T1WSE MRI scan illustrating the transitional slice. This is defined as the first slice in the cartilaginous portion of the SIJ that has a visible portion of the ligamentous joint (arrow) when viewing DICOM images of semicoronal slices through the SIJ from anterior to posterior. T1WSE: T1-weighted spin echo; MRI: magnetic resonance imaging; DICOM: Digital Imaging and Communication in Medicine; SIJ: sacroiliac joints.

a pair of scans (baseline, 2 yrs) per case was estimated. Because reliability for baseline as well as 2-year change scores was at least good [intraclass correlation coefficient (ICC)  $\geq 0.6$ ] for most features (status/change ICC for fat metaplasia = 0.72/0.68, for erosion = 0.60/0.59, for backfill = 0.86/0.55, for ankylosis = 0.98/0.79), further validation was undertaken

after debriefing of discrepant cases and clarification of lesion definitions for scoring purposes; (2) in reading exercise 1, 4 readers scored 45 cases with baseline and 2-year scans randomly selected from the cohort to assess whether interobserver reliability for the 4 readers was consistently evident, particularly for change scores, when assessed blinded to timepoint; (3) in



*Figure 3.* Example of SSS method applied to a T1WSE MRI scan from a patient with SpA assessed at baseline and 2 years. Scoring for ER, BF, FAT, and ANK is indicated on the schematic of the SIJ to the right of each MRI scan. Each lesion is indicated as being present/absent on a dichotomous basis (present = 1, absent = 0). A. SSS scores for ER and BF at baseline and 2 years. B. SSS scores for fat and ANK at baseline and 2 years. Fat metaplasia must extend more than 1 cm in depth from the joint surface to be scored. At 2 years, left upper iliac quadrant is not scored even though fat metaplasia is present because it does not extend more than 1 cm from the joint surface. SSS: Sacroiliac Joint Structural Score; T1WSE: T1-weighted spin-echo; MRI: magnetic resonance imaging; SpA: spondyloarthritis; ER: erosion; BF: backfill; FAT: fat metaplasia; ANK: ankylosis; SIJ: sacroiliac joints.

reading exercise 2, 2 of the 4 readers independently scored an additional 102 cases blinded to timepoint with available baseline and 2-year scans to assess reliability for baseline and 2-year change scores across the spectrum of patients in the cohort; and (4) in reading exercise 3, we assessed whether sufficient reliability of change scores could be attained in as short a time frame as 1 year. Baseline and 1-year scans from 40 cases were assessed blinded to timepoint by 3 of the 4 readers. The selection of these cases required that the baseline STIR MRI scan had evidence of BME, which had been scored previously using the SPARCC SIJ score prior to any development work on the SSS method. The rationale for this selection is that most current trials on nonradiographic axial SpA now require a positive MRI of the SIJ according to the presence of BME as an inclusion criterion.

**Statistics.** We used descriptive statistics to assess the number (percentage) of patients with any change and the mean (SD) change in each of the 4 structural lesion scores over 2 years. Interobserver reliability for baseline, 1-year, and 2-year change scores was assessed using ICC (3,1). A 2-way mixed effects model with patient as a random factor and observer as a fixed factor was used, and the results are given as single measures. The ICC are presented as absolute agreement for the individual reader pairs and for all readers together (overall ICC). An ICC value of < 0.4 was designated fair;  $\geq 0.4$  but < 0.6 moderate;  $\geq 0.6$  but < 0.8 good;  $\geq 0.8$  but < 0.9 very good; and  $\geq 0.9$  excellent. We also calculated the smallest detectable change (SDC), which provides an absolute measure of agreement using the Bland-Altman 80% levels of agreement<sup>20</sup>.

## RESULTS

### *Baseline characteristics and descriptive SPARCC SSS data.*

The majority of patients in all 3 exercises were men with longstanding disease receiving TNF- $\alpha$  inhibitor therapy (Table 1). A minority received only prescribed NSAID therapy and some patients received no prescribed therapy, but took over-the-counter antiinflammatory agents. At baseline, the majority of patients in each of the 3 exercises had at least 1 SIJ quadrant with fat metaplasia, erosion, and backfill in the SIJ, and about half in each exercise had at least 1 SIJ quadrant with ankylosis. Mean SSS erosion score decreased over followup while scores for other structural lesions increased. At the patient level, an increase in SSS fat metaplasia score was observed in a mean of 19.7% and a decrease in 15.6% for the 2 readers who read all 147 available baseline and 2-year scans. Similarly, for SSS erosion score, an increase was observed in a mean of 14.6% patients and a decrease in 37.8%. An increase in SSS backfill was observed in a mean of 23.1% of patients and a decrease in 16.7%. The SSS method detected change in structural lesion score for each lesion as soon as 1 year in the majority of patients assessed in exercise 3.

**Feasibility and interobserver reliability.** The time required to assess a pair of scans per case for all structural lesions ranged from 5–15 min. The selection of 5 slices from the transitional slice anteriorly allowed the assessment of structural lesions in the bone marrow of both sides of the joint cavity in all patients. Interobserver reliability for status score assessed at baseline was good to excellent for ankylosis, consistently good for fat metaplasia, moderate to good for erosion, and fair to good for backfill (Table 2). Very good to excellent reliability was achieved by some reader pairs.

Interobserver reliability for change scores was moderate to good for all structural lesions and for all exercises with the exception of erosion in exercise 1 despite the relatively small changes in scores that were recorded (Table 3). Moderate to good reliability was evident for scoring 1-year change in all structural lesions and was highest for fat metaplasia when both ICC and SDC values were compared. SDC values were consistently less than 5% of the maximum score for fat metaplasia in all 3 exercises. Backfill was the most difficult lesion to detect reliably with SDC values of between 5–10% of the maximum score for each of the 3 exercises.

## DISCUSSION

We have developed and conducted preliminary validation of a scoring method for structural lesions in the SIJ that is based on the same scoring principles used in the SPARCC MRI SIJ inflammation scoring method. The approach to the selection of MRI slices is anatomically defined, the majority of the cartilaginous portion of the joint is assessed on consecutive slices in the semicoronal plane, and scoring is dichotomous (present/absent), which simplifies assessment and improves reliability. Assessment of status and change scores on a pair of scans can be completed in a feasible time frame. Interobserver reliability for status scores is good to excellent for all lesions and can be achieved early in the calibration process. Good reliability can also be achieved for change scores at the 1-year followup.

Scoring with this method begins with the identification of the transitional slice and then proceeds by evaluating 5 consecutive slices ventrally through the cartilaginous portion of the joint. Scoring slices dorsal to the transitional slice is not appropriate because those slices are predominantly or entirely composed of the ligamentous portion of the joint. In addition, scoring slices more ventrally often results in failure to depict the joint clearly. In both cases, these additional ventral or dorsal slices do not provide useful additional information in the context of scoring structural damage, depict areas that are frequently the hardest to interpret even in normal subjects, and including them in the exercise reduced reader reliability (data not shown). A variable that could influence the application of this method is that the comparison of scores between timepoints can be more challenging if the coronal tilt of slices through the SIJ varies between timepoints. This is easily prevented by stipulating appropriate methodological detail in protocols for imaging in clinical trials and research. We recommend the following protocol for defining the tilted coronal plane. After a triplanar series of scout images, a dedicated small field-of-view sagittal scout series of the sacrum is performed. With reference to the midline sagittal image of the sacrum, the tilted coronal sequences are angled parallel to the posterior cortex of the second sacral vertebral body, which also represents the anterior border of the sacral spinal canal at the S2 level.

Table 1. Baseline patient characteristics and SPARCC MRI SSS, and change over 2 years (exercises 1 and 2) or 1 year (exercise 3) from patients with SpA.

Patient Characteristics	Exercise 1, n = 45	Exercise 2, n = 102	Exercise 3, n = 40
No. readers	4	2	3
Age, mean (SD)	41.4 (12.6)	39.9 (11.3)	36 (9.4)
Males, n (%)	34 (76)	77 (75)	34 (85)
Symptom duration, yrs, mean (SD)	15.9 (10.5)	16.9 (9.9)	14 (8.4)
BASDAI, mean (SD)	5.3 (2.5)	4.7 (2.4)	5.5 (2.2)
CRP, mg/l, mean (SD)	13.5 (15.0)	13.0 (17.0)	11.5 (13.7)
Taking only NSAID, n (%)	17 (38)	39 (38)	17 (43)
Taking TNF- $\alpha$ inhibitor, n (%)	23 (51)	56 (55)	21 (53)
SSS fat score, 0–40*, mean (SD) (range)			
Baseline	3.38 (4.07) (0–17)	3.84 (6.54) (0–31.5)	4.61 (6.30) (0–26)
Change	0.21 (1.43) (–4.25 – +5.75)	0.33 (2.03) (–6 – +11)	0.73 (1.80) (–4 – +6)
SSS erosion, 0–40*, mean (SD) (range)			
Baseline	3.09 (3.37) (0–15)	3.03 (4.01) (0–21)	4.63 (2.77) (0–18)
Change	–1.32 (2.04) (–8.25 – +4.25)	–0.94 (2.56) (–12 – +6)	–0.63 (1.67) (–9.33 – +6)
SSS backfill, 0–20*, mean (SD) (range)			
Baseline	3.68 (4.34) (0–17.8)	2.76 (3.68) (0–19)	4.93 (3.09) (0–18)
Change	0.5 (2.55) (–7.25 – +10.75)	0.06 (2.45) (–14.5 – +8)	–0.03 (1.61) (–8.33 – +7.33)
SSS ankylosis, 0–20*, mean (SD) (range)			
Baseline	5.27 (7.96) (0–20)	6.16 (8.02) (0–20)	2.36 (3.86) (0–14.7)
Change	0.15 (1.13) (–2.25 – +6.5)	0.36 (1.73) (–5 – +9)	0.58 (1.39) (–1.33 – +7.67)
No. of patients SSS fat, mean, n (%)			
Baseline > 0	25.3 (56)	44 (43)	24.3 (61)
Change > 0	19.8 (44)	34 (33)	17.3 (43)
Change > SDC	19.8 (44)	25.5 (25)	17.3 (43)
No. of patients SSS erosion, mean, n (%)			
Baseline > 0	26.8 (59)	54 (53)	32.3 (81)
Change > 0	24 (53)	52 (51)	32 (80)
Change > SDC	17.3 (38)	27 (26)	20.3 (51)
No. of patients SSS backfill, mean, n (%)			
Baseline > 0	23.8 (53)	43 (42)	31 (78)
Change > 0	24.3 (54)	36 (35)	27.3 (68)
Change > SDC	18.5 (41)	27 (26)	16.7 (42)
No. of patients SSS ankylosis, mean, n (%)			
Baseline > 0	16.5 (37)	47 (46)	13.7 (34)
Change > 0	11.8 (26)	27.5 (27)	12.7 (32)
Change > SDC	11.8 (26)	27.5 (27)	7 (18)

\* Maximum scoring range for SSS score. Baseline SSS score > 0 denotes at least 1 SIJ quadrant in the 5 coronal slices assessed in the SSS method with the structural lesion. Change SSS score > 0 denotes at least 1 SIJ quadrant of the 5 SIJ slices with a new or resolved structural lesion. SPARCC: the Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; SSS: Sacroiliac joint Structural Score; SpA: spondyloarthritis; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; CRP: C-reactive protein; NSAID: nonsteroidal antiinflammatory drug; TNF- $\alpha$ : tumor necrosis factor- $\alpha$ ; SIJ: sacroiliac joint.

A defining principle of the SPARCC methodology for scoring inflammatory and structural lesions in SpA is the application of a dichotomous scoring method based on the simple presence or absence of a lesion per slice in an anatomically defined location such as an SIJ quadrant. The pathological abnormalities visible on MRI often include mixed lesions with complex appearances that may hinder scoring approaches based on estimates of percent volume of an anatomical region occupied by the lesion, especially in the SIJ. Several lesions may be evident concomitantly even in the same SIJ quadrant, and reliable assessment of the volume of the SIJ quadrant occupied by a specific lesion in a joint as morphologically complex as the SIJ may be exceptionally difficult. For example, a lesion visible only in anterior slices will occupy a greater percentage of an SIJ

quadrant or the iliac/sacral portion of the joint as compared to the same-sized lesion present only in dorsal slices. For lesions extending across several slices from ventral to dorsal, scoring based on the percent volume occupied by the lesion requires that the reader attempt to mentally compile a 3-D estimate based on all slices assessed.

The SDC was calculated based on 80% limits of agreement<sup>20</sup> and resulted in values that are higher than the mean change scores observed in all 3 exercises. This indicates that there is a risk that patients may be misclassified as having progression when in fact there is measurement error. In our dataset, however, it is also important to note that change in structural lesions may occur in both directions, i.e., increase or decrease, and the mean score reflects a summation at the group level. This limits

Table 2. ICC values [overall and mean (range) of all possible reader pairs] for SPARCC MRI SIJ SSS baseline scores for patients with SpA in 3 reading exercises. Mean ICC (range) = mean of ICC for all possible reader pairs with range of ICC scores.

Exercise No.	Readers	Fat Metaplasia		Erosion		Backfill		Ankylosis	
		Overall ICC	Mean ICC (range)	Overall ICC	Mean ICC (range)	Overall ICC	Mean ICC (range)	Overall ICC	Mean ICC (range)
Exercise 1, n = 45	4	0.78	0.73 (0.58–0.86)	0.62	0.55 (0.45–0.77)	0.66	0.63 (0.51–0.73)	0.98	0.98 (0.96–0.99)
Exercise 2, n = 102	2	0.78	—	0.58	—	0.35	—	0.98	—
Exercise 3, n = 40	3	0.71	0.71 (0.57–0.81)	0.65	0.66 (0.60–0.71)	0.66	0.64 (0.51–0.88)	0.79	0.78 (0.72–0.87)

ICC: intraclass correlation coefficient; SPARCC: the Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; SIJ: sacroiliac joint; SSS: Sacroiliac joint Structural Score; SpA: spondyloarthritis.

Table 3. Interobserver ICC values [overall and mean (range) of all possible reader pairs] and SDC for 2-year (exercises 1 and 2) and 1-year (exercise 3) change in SPARCC MRI SIJ SSS in patients with SpA.

Exercise Patient No.	Readers	Fat Metaplasia			Erosion			Backfill			Ankylosis		
		Overall ICC	Mean ICC (range)	SDC	Overall ICC	Mean ICC (range)	SDC	Overall ICC	Mean ICC (range)	SDC	Overall ICC	Mean ICC (range)	SDC
Exercise 1, n = 45	4	0.44	0.44 (0.29–0.64)	0.9 (2.3)	0.39	0.38 (0.27–0.45)	1.4 (3.5)	0.56	0.54 (0.46–0.66)	1.3 (6.6)	0.64	0.65 (0.58–0.74)	0.5 (2.3)
Exercise 2, n = 102	2	0.58	—	1.3 (3.3)	0.40	—	2.1 (5.3)	0.49	—	1.8 (9.0)	0.73	—	0.9 (4.4)
Exercise 3, n = 40	3	0.61	0.61 (0.51–0.79)	0.9 (2.2)	0.54	0.54 (0.48–0.60)	1.7 (4.4)	0.42	0.43 (0.19–0.63)	1.2 (5.8)	0.47	0.48 (0.29–0.69)	1.2 (6.0)

ICC: intraclass correlation coefficient; SPARCC: the Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; SIJ: sacroiliac joint; SSS: Sacroiliac joint Structural Score; SpA: spondyloarthritis; SDC: smallest detectable change.

interpretation and conclusion about the potential for misclassification of measurement error based on evaluation of group mean data.

Several scoring methods have been developed for quantification of fat metaplasia, and the reliability for assessment of status scores was reported as very good although none as yet have been validated to show that change scores can be reliably detected<sup>6,9,10,11,14</sup>. One report assessed fat lesions dichotomously (present/absent) according to SIJ quadrants and a 0–8 scoring range based on a global evaluation of the SIJ rather than scoring of SIJ quadrants in individual slices<sup>6</sup>. Development of new lesions correlated with resolution of inflammation and a significant difference was noted between patients receiving etanercept versus those taking sulfasalazine as soon as 24 weeks. A second method grades fat metaplasia in both cartilaginous and ligamentous compartments and grades severity according to the extent of subcortical bone affected (0 = no fat, 1 = < 25%, 2 = 25–50%, 3 = > 50%)<sup>14</sup>. A weighting of 1 is added for fat metaplasia extending  $\geq 1$  cm beneath the joint surface. Fat metaplasia may be distributed in a very heterogeneous manner in the iliac and sacral bones, and is typically located adjacent to the subchondral bone where it may form a linear band of variably increased signal on the T1-weighted scan<sup>21</sup>. Our preliminary development work showed that reliable detection of fat metaplasia can be enhanced by scoring lesions that extend at least 1 cm in depth from the joint surface without affecting responsiveness.

Erosion in the SIJ has been quantified according to a

grading scheme based on the number of erosions per SIJ (1 = 1–2, 2 = 3–5, 3 = > 5 erosions per SIJ)<sup>6</sup>. Reliability for status score was very good, but data for change scores was not reported. In a second method, erosion was graded in both the cartilaginous and ligamentous compartments, and severity graded according to the extent of subcortical bone affected (0 = no erosion, 1 = < 25%, 2 = 25–50%, 3 = > 50%)<sup>14</sup>. Reliability for status score was good, but data for change score was not reported. A limitation of scoring based on the number of erosions per SIJ is that it is not unusual to observe erosion affecting the entire vertical height of the iliac or sacral bone on a coronal scan so that one cannot consider these as discrete lesions. In addition, this scheme appears to preclude the possibility that there may be a substantial reduction in the size of an erosion without affecting the number of erosions.

The scoring methodology we have developed suggests that reliable detection of change in structural lesions in the SIJ may be possible even after 1 year. Ankylosis is defined by the presence of bright marrow signal traversing the joint space between the iliac and sacral bones and is therefore easier to discern reliably on T1-weighted MRI than erosion, which requires loss of dark signal, signifying breach of cortical bone. Nevertheless, reliability for detecting change in erosion and ankylosis was comparable, suggesting that application of validated and standardized lesion definitions and calibration based on DICOM images illustrating the spectrum of abnormalities and changes over time can be an effective mode of knowledge transfer. Backfill presents a

more heterogeneous MRI appearance with increased signal on the T1-weighted scan that is clearly demarcated from adjacent normal marrow by irregular dark signal reflecting sclerosis at the border of the eroded bone. This is more difficult to define in a standardized manner and is the likely reason why more extensive calibration is required to achieve more reliable detection of this lesion. We have shown that resolution of inflammation and reduction in erosion is significantly associated with the development of backfill<sup>22</sup>, although further validation of this lesion using computed tomography would be helpful.

We have developed and validated a scoring methodology for the assessment of structural lesions on MRI in the SIJ of patients with SpA that is based on the same principles as the widely used SPARCC MRI SIJ inflammation score. These include selection of MRI slices according to well-defined anatomical principles, the scoring of lesions on a dichotomous basis as being present/absent on consecutive, oblique coronal slices through the cartilaginous portion of the joint, and calibration of readers using standardized definitions and DICOM-based reference cases. Further validation should now be undertaken aimed at assessment of responsiveness and the pathophysiological and prognostic significance of these lesions observed on MRI.

## REFERENCES

- Maksymowych WP, Morency N, Conner-Spady B, Lambert RG. Suppression of inflammation and effects on new bone formation in ankylosing spondylitis: evidence for a window of opportunity in disease modification. *Ann Rheum Dis* 2013;72:23-8.
- Haroon N, Inman RD, Learch TJ, Weisman MH, Lee M, Rahbar MH, et al. The impact of tumor necrosis factor alpha inhibitors on radiographic progression in ankylosing spondylitis. *Arthritis Rheum* 2013;65:2645-54.
- Maksymowych WP, Chiowchanwisawakit P, Clare T, Pedersen SJ, Ostergaard M, Lambert RG. Inflammatory lesions of the spine on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis: evidence of a relationship between inflammation and new bone formation. *Arthritis Rheum* 2009;60:93-102.
- Weber U, Østergaard M, Lambert RG, Maksymowych WP. The impact of MRI on the clinical management of inflammatory arthritides. *Skeletal Radiol* 2011;40:1153-73.
- Chiowchanwisawakit P, Lambert RG, Conner-Spady B, Maksymowych WP. Focal fat lesions at vertebral corners on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis. *Arthritis Rheum* 2011;63:2215-25.
- Song I-H, Hermann KG, Haibel H, Althoff CE, Poddubny D, Listing J, et al. Relationship between active inflammatory lesions in the spine and sacroiliac joints and new development of chronic lesions on whole-body MRI in early axial spondyloarthritis: results of the ESTHER trial at week 48. *Ann Rheum Dis* 2011;70:1257-63.
- Hermann KG, Baraliakos X, van der Heijde D, Jurik AG, Landewé R, Marzo-Ortega H, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial SpA. *Ann Rheum Dis* 2012;71:1278-88.
- Weber U, Pedersen SJ, Ostergaard M, Rufibach K, Lambert R, Maksymowych WP. 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.
- Baraliakos X, van der Heijde D, Braun J, Landewé RB. 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.
- Weber U, Lambert RG, Østergaard M, Hodler J, Pedersen SJ, Maksymowych WP. The diagnostic utility of MRI in spondyloarthritis: an international multicentre evaluation of 187 subjects (The MORPHO study). *Arthritis Rheum* 2010;62:3048-58.
- Madsen KB, Schiøtz-Christensen B, Jurik AG. Prognostic significance of magnetic resonance imaging changes of the sacroiliac joints in spondyloarthritis — a followup study. *J Rheumatol* 2010;37:1718-27.
- Marzo-Ortega H, McGonagle D, O'Connor P, Hensor EM, Bennett AN, Green MJ, et al. Baseline and 1-year magnetic resonance imaging of the sacroiliac joint and lumbar spine in very early inflammatory back pain. Relationship between symptoms, HLA-B27 and disease extent and persistence. *Ann Rheum Dis* 2009;68:1721-7.
- Wick MC, Weiss RJ, Jäschke W, Klauser AS. Erosions are the most relevant magnetic resonance imaging features in quantification of sacroiliac joints in ankylosing spondylitis. *J Rheumatol* 2010;37:622-7.
- Madsen KB, Jurik AG. Magnetic resonance imaging grading system for active and chronic spondyloarthritis changes in the sacroiliac joint. *Arthritis Care Research* 2010;62:11-18.
- Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998;25:198-9.
- Wanders A, Heijde D, Landewé R, Béhier JM, Calin A, Olivieri I, et al. Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. *Arthritis Rheum* 2005;52:1756-65.
- Gladman DD, Rahman P, Cook RJ, Shen H, Zummer M, Thomson G, et al. The Spondyloarthritis Research Consortium of Canada Registry for Spondyloarthritis. *J Rheumatol* 2011;38:1343-8.
- Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. *Arthritis Rheum* 2005;53:703-9.
- Østergaard M, Maksymowych WP, Pedersen SJ, Chiowchanwisawakit P, Lambert RG. Structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis – definitions, assessment system, and reference image set. *J Rheumatol Suppl.* 2009 Dec;84:18-34.
- Navarro-Compán V, van der Heijde D, Ahmad HA, Miller CG, Wolterbeek R, Landewé R. Measurement error in the assessment of radiographic progression in rheumatoid arthritis (RA) clinical trials: the smallest detectable change (SDC) revisited. *Ann Rheum Dis* 2014;73:1067-70.
- Weber U, Pedersen SJ, Zubler V, Rufibach K, Chan SM, Lambert RG, 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.
- Maksymowych WP, Wichuk S, Chiowchanwisawakit P, Lambert RG, Pedersen SJ. Fat metaplasia and backfill are key intermediaries in the development of sacroiliac joint ankylosis in patients with ankylosing spondylitis. *Arthritis Rheumatol* 2014 Jul 21 (E-pub ahead of print).