

# Spondyloarthritis Research Consortium of Canada Scoring System for Sacroiliitis in Juvenile Spondyloarthritis / Entesitis-related Arthritis: A Reliability, Validity, and Responsiveness Study

Jyoti Panwar , Shirley M.L. Tse, Lillian Lim, Mirkamal A. Tolend, Shilpa Radhakrishnan, Muayad Salman, Rahim Moineddin , Andrea S. Doria, and Jennifer Stimec

**ABSTRACT.** *Objective.* Intra- and interreader reliability, construct validity, and responsiveness of the Spondyloarthritis Research Consortium of Canada (SPARCC) magnetic resonance imaging (MRI) scoring system were investigated for scoring sacroiliitis in patients with juvenile spondyloarthritis (JSpA)/entesitis-related arthritis (ERA) who have received biologic and/or nonbiologic treatment. *Methods.* Ninety whole-body MRI examinations with dedicated oblique coronal planes of the sacroiliac joints in 46 patients were independently reviewed and scored by 2 pediatric musculoskeletal radiologists, blinded to clinical details, using the SPARCC system. Intra- and interreader reliability was assessed by intraclass correlation coefficients (ICC). Construct validity testing was done by (1) correlating the SPARCC MRI scores of sacroiliitis with clinical disease activity indicators (cross-sectional validity), and (2) correlating the change in the MRI score with the change in clinical indicators before and after treatment (longitudinal validity). Responsiveness of the MRI and clinical indicators was also evaluated, grouped by biologic and nonbiologic treatment. *Results.* When applied in children with JSpA/ERA, the SPARCC showed almost perfect intra- and interreader reliability (ICC 0.79–1.00). There was poor cross-sectional and longitudinal correlation between clinical assessment indicators and MRI scoring. SPARCC scores showed higher responsiveness to treatment-related change than most clinical outcome measures. Three clinical outcome measures correlated longitudinally with SPARCC score in nonbiologic treatment: active joint count ( $r = 0.72$ ,  $p < 0.001$ ), FABER (Flexion, Abduction, External Rotation) test ( $r = 0.58$ ,  $p = 0.012$ ), and physician's global assessment ( $r = 0.61$ ,  $p = 0.034$ ). *Conclusion.* The SPARCC MRI scoring system is a reliable tool with relatively higher responsiveness than clinical indicators and is suitable for objective quantification of sacroiliitis when applied to pediatric patients with JSpA/ERA. (J Rheumatol First Release February 1 2019; doi:10.3899/jrheum.180222)

## Key Indexing Terms:

SPONDYLOARTHRITIS RESEARCH CONSORTIUM OF CANADA  
MAGNETIC RESONANCE IMAGING  
JUVENILE SPONDYLOARTHRITIS

SACROILITIS  
SCORING SYSTEM  
ENTHESITIS-RELATED ARTHRITIS

From the Joint Department of Medical Imaging, University Health Network, University of Toronto; Department of Diagnostic Imaging, The Hospital for Sick Children, University of Toronto; Division of Pediatric Rheumatology, The Hospital for Sick Children, University of Toronto, Toronto, Ontario; Division of Pediatric Rheumatology, Department of Pediatrics, The University of Alberta, Edmonton, Alberta; Kanchi Kamakoti CHILDS Trust Hospital, Chennai; Department of Radiology, Christian Medical College, Vellore, India; Makassed Hospital, Jerusalem, Israel; Department of Family and Community Medicine, University of Toronto, Toronto, Ontario, Canada.

J. Panwar, MD, FRCR, Joint Department of Medical Imaging, University Health Network, University of Toronto, and Department of Diagnostic Imaging, The Hospital for Sick Children, University of Toronto, and Department of Radiology, Christian Medical College; S.M. Tse, MD, Division of Pediatric Rheumatology, The Hospital for Sick Children, University of Toronto; L. Lim, MD, FRCPC, Division of Pediatric Rheumatology, Department of Pediatrics, The University of Alberta; M.A. Tolend, Department of Diagnostic Imaging, The Hospital for Sick Children, University of Toronto; S. Radhakrishnan, MD, Kanchi Kamakoti CHILDS Trust Hospital; M. Salman, MD, Makassed Hospital;

R. Moineddin, PhD, Department of Family and Community Medicine, University of Toronto; A.S. Doria, MD, PhD, MSc, Department of Diagnostic Imaging, The Hospital for Sick Children, University of Toronto; J. Stimec, MD, Department of Diagnostic Imaging, The Hospital for Sick Children, University of Toronto.

Address correspondence to Dr. J. Panwar, Department of Diagnostic Imaging, 555 University Ave., Toronto, Ontario M5G 1X8, Canada; or Department of Radiology, Christian Medical College, Vellore-632004, India. E-mail: drjyotcmch@gmail.com, Jyoti.panwar@sickkids.ca

Accepted for publication September 12, 2018.

Juvenile spondyloarthritis (JSpA) affects the peripheral and axial skeleton and has a strong genetic association with HLA-B27. It is most commonly referred to as entesitis-related arthritis (ERA), a subtype of juvenile idiopathic arthritis (JIA) as defined by the International League of Associations for Rheumatology (ILAR) classification for

childhood arthritis<sup>1</sup>. ERA accounts for about 15–20% of JIA cases<sup>2,3,4</sup>, with a peak age of disease onset at 11.7 years<sup>5</sup> and predilection for male sex<sup>6</sup>. JSpA/ERA often have a positive family history for HLA-B27-associated disease including reactive arthritis, ankylosing spondylitis (AS), inflammatory bowel disease, psoriasis or psoriatic arthritis, and acute iritis or uveitis<sup>7</sup>.

Peripheral arthritis and enthesitis usually affect the lower extremities and are the most common presentations in JSpA/ERA<sup>8</sup>. Axial involvement of the sacroiliac (SI) joints and spine is uncommon at disease onset<sup>9,10</sup> but may develop within 5 to 10 years in up to 40% of JSpA/ERA patients<sup>6,11</sup> and is associated with higher morbidity, pain scores, and progression to AS<sup>12,13</sup>. Clinical symptoms of axial involvement are nonspecific but can include lumbar or buttock pain. Clinical sacroiliitis is elicited by pain with direct palpation or through provocative maneuvers, such as the FABER (Flexion, Abduction, External Rotation) or Gaenslen's tests. Magnetic resonance imaging (MRI) is far more sensitive than radiographs in the detection of SI inflammation and structural damage<sup>14,15,16</sup>. Early detection of sacroiliitis is important for diagnosis and institution of appropriate management, especially with biologic therapies [anti-tumor necrosis factor (TNF) agents], to ensure improved patient outcomes and quality of life<sup>17</sup>.

MRI offers a more comprehensive evaluation of anatomical structures in the SI joints and spine, along with early detection of asymptomatic inflammatory lesions, and is unattainable with physical examination alone. This has been demonstrated in patients with JSpA/ERA and has been helpful in providing evidence of sacroiliitis in adolescents with a negative history of back or buttock pain and a normal axial physical examination<sup>18</sup>. Various MRI scoring methods exist to quantitatively evaluate the inflammation and bone marrow edema (BME) in patients with sacroiliitis<sup>19,20,21,22</sup>. These scoring systems differ regarding the MRI planes and sequences used to detect inflammation, a more precise quantification for unit of interest (SI joint divided into halves or quadrants), number of slices used to score, global versus more extensive grading, and the site of the inflammatory lesion to be scored<sup>23</sup>. Among these, the Spondyloarthritis Research Consortium of Canada (SPARCC) scoring system is the most advantageous because it is both comprehensive and objectively scored according to a standardized measurement protocol<sup>21</sup>. While other scoring systems account only for the presence or absence of BME, the SPARCC scoring system incorporates 2 other MRI indices of potential clinical significance: signal intensity and the 3-D extent of inflammation<sup>24</sup>.

The validity and reliability of the SPARCC scoring system has been shown in adult SpA and AS populations with sacroiliitis, and thus is one of the most accepted and used worldwide. Unfortunately, no similar studies have yet been conducted in the pediatric population to provide objective

measurements for diagnostic purposes and disease activity monitoring. In our study, we investigated the intra- and inter-reader reliability of the SPARCC scoring system in children and adolescents with sacroiliitis in the setting of JSpA/ERA. This study also included exploratory analysis to determine trends in correlation between the SPARCC scores and clinical variables of disease activity.

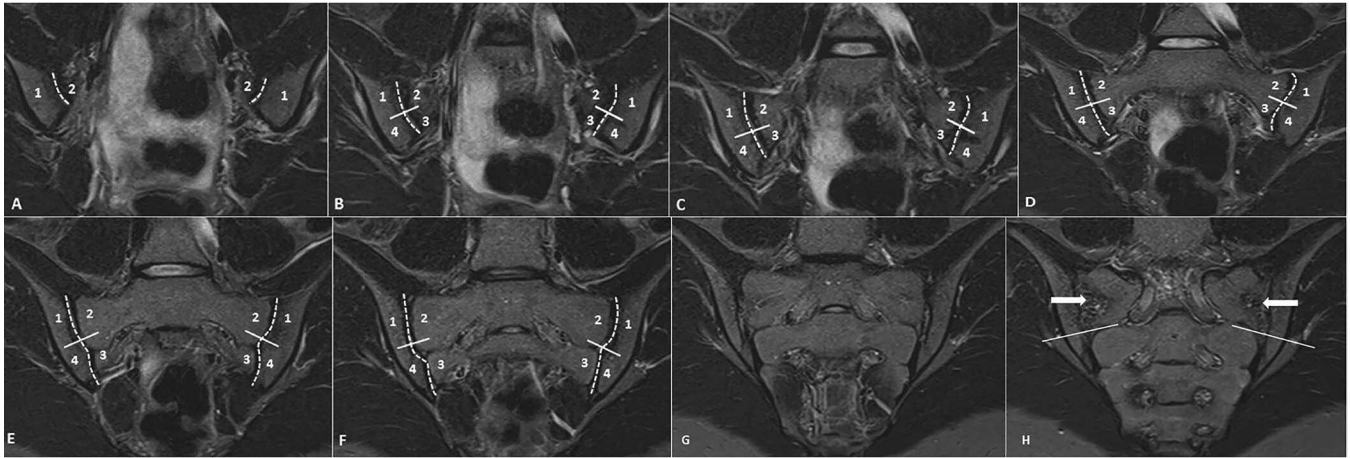
## MATERIALS AND METHODS

This cross-sectional retrospective study was undertaken at a large pediatric quaternary referral center and approved by the institutional research ethics board (no. 1000044486). The consent from all patients was waived owing to the retrospective features of the study.

*Patient selection.* All patients included were under the age of 16 years, had a confirmed or suspected clinical diagnosis of JSpA (European Spondyloarthritis Study Group criteria) and/or ERA (ILAR criteria), and had a whole-body (WB) MRI performed between January 2008 and December 2016. Electronic medical records identified 70 eligible patients. Of these, 24 patients were excluded because of inadequate or incomplete MR examinations attributed to motion artifacts or absence of dedicated oblique coronal of SI joints necessary for SPARCC scoring evaluation. A total of 90 MRI studies were analyzed for the 46 patients included in the study; 15 patients underwent more than 1 MRI study while receiving treatment. Demographic, clinical, and laboratory data were also collected for the 46 patients enrolled. For 15 patients, pretreatment and posttreatment clinical and laboratory data were collected.

*MRI protocol.* All MRI studies were done as part of a WB-MRI protocol that we use for the evaluation of JSpA/ERA. All examinations were performed on a 1.5-T MRI system (Magnetom Avanto, Siemens), using a dedicated multichannel surface coil system with the patient supine. Integrated head, neck, spine, body, and peripheral angiography surface coils were used for contiguous scanning. Images were acquired at multiple stations with the patient free-breathing and were subsequently reconstructed using the vendor-specific software package (Siemens Composing, Siemens) to form a WB image. All examinations included oblique coronal short-tau inversion recovery (STIR) imaging (repetition time msec/echo time msec of 2250/69 with a field of view of 25 cm) of the SI joints as a part of the WB-MRI. Scans were obtained in an oblique coronal plane parallel to the long axis of the sacral vertebrae (S1–S2) using 4 mm of slice thickness with a slice gap of 4.5 mm.

*MRI analysis.* The SPARCC scoring system quantifies BME within the iliac and sacral bones along the SI joint by using the corresponding increased MRI signal detected on STIR images. Quantification is performed on 6 consecutive oblique coronal slices through the SI joint, encompassing most of the cartilaginous and synovial portion of the joint that have at least 1 cm of visible vertical height. Each SI joint is divided into iliac and sacral portions. When the vertical height of the SI joint measures > 3 cm, each half is further divided into 2 quadrants, upper and lower (Figure 1). As a result, a maximum of 48 quadrants of SI joints are created within 6 MRI slices. This rule is applied to the anterior and posterior aspects of the SI joints. However, at the posterior aspect, the joint is naturally divided into upper and lower quadrants by intervening fat and fibrous tissue (Figure 1). Hence BME within the ligamentous portion was not scored (Figure 2 and Figure 3). Three MRI indices were evaluated. First, the increased STIR signal (BME) was evaluated: the presence of bone marrow edema was identified based on comparison with signal from the center of the sacrum at the same cranio-caudal level, because this site is less prone to fatty change or inflammation (Figure 2). If the center of the sacrum was not visible on image, then the adjacent section that provided a reference for comparison at that level was used. One point was granted per quadrant with increased STIR signal, for a maximum score of 48. Second were lesions exhibiting intense signal (intense edema); the adjacent presacral vein signal was used as a reference (Figure 2). One point was granted per SI joint per slice for a maximum score of 12.



**Figure 1.** A–F. Contiguous oblique coronal STIR images depict the zonal division of SI joints in a patient with a history of sacroiliitis. The initial 6 images through most of the cartilaginous and synovial portion of the joint, from anterior to posterior, are analyzed. The SI joints of the most anterior slice (start slice) A [vertical height measures > 1 cm (right: 1.8 and left: 1.7 cm)] is divided into 2 portions: (1) iliac and (2) sacral. The vertical height of the SI joints measuring > 3 cm (B–F) are further divided into upper and lower quadrants, creating a total of 4 quadrants per joint: (1) upper iliac, (2) upper sacral, (3) lower sacral, and (4) lower iliac. The posterior-most slice (end slice; F) is the sixth slice and includes the synovial portion of the joint at the inferior aspect. G–H. Further posterior slices show natural divisions of the joint into the upper (ligamentous portion) and lower (synovial portion) quadrants by intervening fat and fibrous tissue (arrows in H). The bone marrow edema along the ligamentous portion was not scored if this posterior aspect of the SI joint was included in the 6 contiguous slices. STIR: short-tau inversion recovery; SI: sacroiliac.



**Figure 2.** A. Oblique coronal STIR image shows deep edema as a contiguous increased signal extends over 1 cm from the articular surface on either side of the right SI joint measured perpendicular to the vertical axis of the joint (double-head arrows). The center of the sacrum at the same craniocaudal level (star) was used as a reference for normal when assessing bone marrow signal adjacent to the joint. B. Oblique coronal STIR image in a different patient with MR evidence of sacroiliitis demonstrates intense edema [signal equal to that from the presacral vein (black arrow)] along the iliac side (white thick arrows) of both the SI joints. Note the less-intense high signal (thin arrow) along the right iliac articular surface. C. Oblique coronal STIR image through the posterior aspect of the joint in a different patient depicts bone marrow signal changes along the iliac side of the ligamentous compartment (upper quadrants of the joint) divided naturally by intervening fat and fibrous tissue (arrows). The bone marrow edema along the ligamentous portion of the joint (along the upper iliac quadrants) has not been scored. STIR: short-tau inversion recovery; SI: sacroiliac; MR: magnetic resonance.

Finally, lesion depth (extensive edema) was evaluated — a lesion demonstrating continuous high signal and presenting with a depth of 1 cm or greater from the articular surface measured perpendicular to the vertical axis of the joint (Figure 2). One point was granted per SI joint per slice for a maximum score of 12. The maximum total score was 72 as demonstrated in the scoring sheet (Supplementary Figure 1, available with the online version of this article)<sup>21,23</sup>.

The MRI studies were independently scored using the SPARCC system by 2 pediatric musculoskeletal radiologists, who were blinded to the clinical information of the study patients. Each radiologist performed the task twice with a 6-week interval between the first and second round of readings. A single SPARCC score of 1 was taken as nonspecific BME. Only a score of 2 or greater was considered to have radiologic evidence of sacroiliitis (Figure 3).

**Clinical data acquisition.** There are no clinical diagnostic scoring systems

or clinical maneuvers that are highly sensitive and specific to represent sacroiliitis. However, in addition to palpating SI joint tenderness, focused physical examination measures that target clinical assessment of the SI joints include the Schober and FABER tests<sup>25,26,27</sup>. Further, surrogate assessments of global disease activity in JSpA/ERA can be obtained by 2 commonly used methods. The first method is the physician's global assessment of disease activity (PGA), which represents both peripheral and axial arthritis. It uses a visual analog scale (VAS) with a range from 0 to 10, where 0 is no disease activity and 10 reflects severe disease activity. The VAS is filled out and based on the expert clinical opinion of the physician after considering all aspects of the patient's arthritis including joint symptoms, physical examination, laboratory investigations, physical function questionnaires, and imaging as available. The second method is the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) method, which is a validated patient





internal rotation by intermalleolar distance), and SpA questionnaires [BASDAI and the Bath Ankylosing Spondylitis Functional Index (BASFI)]. *Statistical analysis.* One-way random effects, absolute agreement, and single-rater version of the intraclass correlation coefficients (ICC) were calculated to assess both intra- and interobserver agreement in SPARCC radiologic scoring of sacroiliitis. Spearman correlation coefficients were used to assess the cross-sectional and longitudinal change score agreement between radiologic (i.e., SPARCC) and clinical (i.e., PGA and BASDAI) scores of disease activity. The ICC and Spearman correlation coefficients agreement was interpreted in accordance with common guidelines:  $\leq 0.40$  indicating poor,  $> 0.40$  to  $\leq 0.60$  moderate,  $> 0.60$  to  $\leq 0.80$  substantial, and  $> 0.80$  excellent<sup>28,29</sup>. Responsiveness of patients to treatment through SPARCC and clinical outcome measures was calculated by standardized response mean (SRM). SRM is the ratio of observed change and the SD reflecting the variability of the change values (SRM = mean of differences between the baseline and followup values/SD of these differences)<sup>30</sup>. Change scores were calculated by subtracting final score from initial score to produce positive treatment-related change, as per convention. All analyses (descriptive, ICC, and Spearman) were performed using SAS 9.4 (SAS Institute).

## RESULTS

*Study sample.* Forty-six patients with JSpA, 89.1% of whom had ERA, were included in our study (Supplementary Table 1, available with the online version of this article). The patients' mean age at disease onset was 12.7 years; 47.8% of patients were positive for HLA-B27. MRI evidence of sacroiliitis defined as a score of  $\geq 2$  was noted in 24 of the 46 (52.2%) baseline (46 = first available MRI) MRI studies analyzed.

*Reliability of the SPARCC system.* Radiologic SPARCC scoring of sacroiliitis (overall and MR index-specific scores) demonstrated almost perfect agreement (as interpreted by ICC) within and between readers (results summarized in Table 1). Inter- and intrareader reliability of the intensity of

SI joint marrow signal was lower than those of presence and extent of BME. The range of item scores given for the presence of BME was much larger than the range of item scores given for presence of intense edema and extent of edema, hence these were not comparable.

*Construct validity between SPARCC and clinical scores.* Construct validity is the ability of a scale to determine the extent to which a particular measure relates to other measures in a manner that is consistent with theoretically derived hypotheses about the constructs being measured<sup>31</sup>. In the retrospective chart review, there were no definitive clinical outcome measures that could confirm a diagnosis of sacroiliitis, hence an exploratory analysis was performed to test for correlations between the SPARCC scores and several surrogate clinical outcome measures used for sacroiliitis. Cross-sectional single timepoint assessment showed that for targeted binary outcome measures of sacroiliitis including the FABER and Schober tests as well as the SI joint tenderness, SPARCC scores were significantly different depending on the presence or absence of these outcomes (Table 2). None of the clinical measures other than hip abduction measurement and C-reactive protein correlated significantly with SPARCC MRI scores (Table 2). Other tested clinical outcome measures including active joint count, SJC, PGA, BASDAI score, BASFI score, and Childhood Health Assessment Questionnaire (CHAQ) score did not correlate with SPARCC MRI scores, as shown in Table 2.

*Responsiveness of patients' joints to treatment as measured by SPARCC scores and clinical outcome measures.* Two or more treatment-interval MRI examinations of the SI joint were available in 15 of the 46 patients (32.6%) included in the study, of whom 5 were treated only with biologic agents

Table 1. Distribution of SPARCC MR scores, the intra- and interreader reliability assessed by ICC, and responsiveness of scores before and after biologic or nonbiologic treatments, indicated by standardized response mean (SRM).

| Variables   | SPARCC Total Score | SPARCC Scoring System Indices |                           |                          |
|---|--------------------|-------------------------------|---------------------------|--------------------------|
|   |                    | Bone Marrow Edema             | Intensity (Intense Edema) | Extent (Extensive Edema) |
| Average score in all baseline examinations ( $\pm$ SD)                  | 9.2 (12.1)         | 7.6 (9.6)                     | 0.6 (1.2)                 | 1.0 (1.8)                |
| Median score in all baseline examinations (IQR)                         | 2 (0–19)           | 2 (0–16)                      | 0 (0–1)                   | 0 (0–2)                  |
| Average score in baseline examinations showing sacroiliitis ( $\pm$ SD) | 17.8 (11.4)        | 14.7 (8.7)                    | 1.2 (1.4)                 | 1.9 (2.1)                |
| Median score in baseline examinations showing sacroiliitis (IQR)        | 19 (6–24.5)        | 16 (6–20)                     | 1 (0–2)                   | 2 (0–2.5)                |
| SPARCC intrareader reliability (n = 90 examinations)                    |                    |                               |                           |                          |
| Reader 1  | 0.99 (0.98–0.99)   | ICC (95% CI)                  |                           |                          |
| Reader 2  | 1                  | 1                             | 0.79 (0.70–0.86)          | 0.95 (0.93–0.97)         |
| SPARCC interreader reliability  | 0.99 (0.98–0.99)   | 1                             | 0.85 (0.78–0.90)          | 0.97 (0.95–0.98)         |
| SPARCC responsiveness (n = 17, 18 examination pairs)                    |                    |                               |                           |                          |
| SRM in biologic and nonbiologic treatments                              |                    |                               |                           |                          |
| Reader 1 first read   | 0.591, 0.633       | 0.630, 0.634                  | 0.270, 0.413              | 0.429, 0.441             |
| Reader 2 second read  | 0.560, 0.657       | 0.597, 0.648                  | 0.236, 0.517              | 0.307, 0.453             |
| Reader 2 first read   | 0.602, 0.647       | 0.626, 0.657                  | 0.397, 0.473              | 0.443, 0.441             |
| Reader 1 second read  | 0.529, 0.674       | 0.569, 0.667                  | 0.275, 0.488              | 0.322, 0.528             |
| Pooled average of 4 readings  | 0.582, 0.667       | 0.619, 0.666                  | 0.302, 0.481              | 0.384, 0.476             |
| Change score reliability (ICC)  | 0.99, 0.97         | 1.00, 0.98                    | 0.65, 0.81                | 0.92, 0.95               |

Presence of sacroiliitis was defined as SPARCC score of  $\geq 2$ . SPARCC: Spondyloarthritis Research Consortium of Canada; MR: magnetic resonance; ICC: intraclass correlation coefficients; IQR: interquartile range.

Table 2. Cross-sectional correlation between SPARCC MRI scores and various clinical outcome measures.

| Clinical Measures        | No. Cases Analyzed | Mean Spearman Coefficient       | Minimum Spearman Coefficient               | Maximum Spearman Coefficient               |
|--------------------------|--------------------|---------------------------------|--|--|
| Active joint count       | 45                 | 0.088                           | 0.081                                      | 0.097                                      |
| Swollen joint            | 45                 | 0.084                           | 0.072                                      | 0.094                                      |
| PGA                      | 38                 | -0.059                          | -0.070                                     | -0.051                                     |
| Parent's assessment      | 15                 | -0.166                          | -0.180                                     | -0.143                                     |
| Enthesitis count         | 44                 | -0.082                          | -0.108                                     | -0.055                                     |
| Hip abduction            | 25                 | -0.343                          | -0.353                                     | -0.335                                     |
| ESR                      | 37                 | 0.111                           | 0.101                                      | 0.133                                      |
| CRP                      | 37                 | 0.072                           | 0.054                                      | 0.107                                      |
| BASDAI                   | 29                 | -0.013                          | -0.022                                     | -0.002                                     |
| BASFI                    | 26                 | 0.061                           | 0.056                                      | 0.067                                      |
| CHAQ                     | 18                 | -0.039                          | -0.039                                     | -0.039                                     |
| Schober                  | 44                 | 0.575*                          | 0.563*                                     | 0.583*                                     |
| FABER                    | 44                 | 0.255                           | 0.251                                      | 0.258                                      |
| SI joint tenderness      | 44                 | 0.290                           | 0.281                                      | 0.298                                      |
| Binary clinical measures | No. cases analyzed | Average MRI score (+ vs - test) | Minimum MRI score difference (+ vs - test) | Maximum MRI score difference (+ vs - test) |
| FABER                    | 44 (10+, 34-)      | 17.4 vs 6.0                     | 16 vs 6.1                                  | 18.3 vs 6.0                                |
| SI joint tenderness      | 44 (14+, 30-)      | 15.0 vs 5.6                     | 14.2 vs 5.6                                | 15.6 vs 5.6                                |
| Schober                  | 44 (6+, 38-)       | 28.2 vs 5.5*                    | 26.3 vs 5.5*                               | 29.5 vs 5.5*                               |

Spearman correlation coefficients for continuous or ordinal clinical measures; differences in the mean SPARCC scores are given for binary clinical outcome measures. The mean and range of results are from 4 independent readings (2 readings per 2 radiologists). Only hip abduction and ESR measures remained significantly correlated with SPARCC MRI scores. Statistically significant differences in SPARCC MRI scores were observed between binary clinical outcomes. \*  $P < 0.001$  after Bonferroni adjustment. SPARCC: Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; PGA: physician's global assessment; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; CHAQ: Childhood Health Assessment Questionnaire; FABER: Flexion, ABduction, and External Rotation; SI: sacroiliac.

(TNF- $\alpha$  inhibitors), 6 only with nonbiologics (naproxen, methotrexate, sulfasalazine, etc.), and 4 with both nonbiologic and biologic therapies. In total, up to 17 intervals of nonbiologic treatments and 18 intervals of biologic treatments were analyzed pairwise, excluding the cases with missing clinical outcome measures. The SPARCC MRI scores could indicate greater response of patients' joints to treatment than could most clinical outcome measures, including active joint count, SJC, Schober and FABER tests, SI joint tenderness, and PGA scores (as shown in Table 3). Hip abduction showed similar or slightly better responsiveness (by SRM) than did the SPARCC MRI scoring system in both treatment cohorts, but the correlation change scores in these 2 measures were nonexistent. Enthesitis count, BASDAI, and BASFI showed varying SRM compared to SPARCC MRI score by type of treatment, also with no correlation of change. CHAQ, parent's assessment, and laboratory measures were not available close to MRI in the majority of cases and therefore could not be assessed. There were no significant correlations between the change in MRI scores versus the change in clinical measures, except for active joint count ( $r = 0.72$ ,  $p < 0.01$ ) and the FABER test ( $r = 0.58$ ,  $p = 0.01$ ) in the nonbiologic group, suggesting suboptimal

MRI-to-clinical correlation in the magnitude and/or direction of change. Among the 3 indices, the BME index generally showed higher responsiveness compared to the intensity and extent indices (Table 1).

For the 15 selected patients with followup MRI scans, there was a median period of 10.3 months (0.0–24.8 mos) from the time of baseline MRI to the start of biologic treatment, and a median period of 10.3 months (3.7–14.3 mos) from the time of MRI to the time of biologic treatment change. Clinical data about biologic treatment dates were not available for 1 patient who was followed at a local community clinic. For nonbiologic treatments, there was a median period of 14.8 months (0.0–78.1 mos) from the time of baseline MRI to the start of nonbiologic treatment, and a median period of 22.6 months (15.5–37.0 mos) from the time of MRI to the time of nonbiologic treatment change. Because of the limited sample size, we cannot conclude any differences between the biologics and nonbiologic treatment groups.

## DISCUSSION

The SPARCC scoring method was chosen for the evaluation of sacroiliitis in our study, because it has been validated for



Table 3. Longitudinal change observed in the SPARCC scores compared to the change in various clinical measures.

| Items               | Nonbiologic Group |                    |           |  |      | Biologic Group  |                    |           |  |      |
|---------------------|-------------------|--------------------|-----------|--|------|-----------------|--------------------|-----------|--|------|
|                     | Total MRI Score   | SRM Clinical Score | No. Cases | Pearson's Correlation of Change Scores | p    | Total MRI Score | SRM Clinical Score | No. cases | Pearson's Correlation of Change Scores | p    |
| All cases           | 0.67              | –                  | 17        | –                                      | –    | 0.58            | –                  | 18        | –                                      | –    |
| Active joint count  | 0.67              | 0.40               | 17        | 0.72                                   | 0.00 | 0.58            | 0.45               | 18        | 0.22                                   | 0.37 |
| Swollen joint count | 0.67              | 0.50               | 17        | 0.41                                   | 0.10 | 0.58            | 0.36               | 18        | 0.07                                   | 0.77 |
| Enthesitis count    | 0.67              | 0.74               | 17        | 0.42                                   | 0.09 | 0.58            | 0.37               | 18        | 0.02                                   | 0.93 |
| Schober             | 0.67              | 0.24               | 17        | 0.36                                   | 0.16 | 0.58            | 0.00               | 18        | 0.10                                   | 0.69 |
| Hip abduction       | 1.13              | –1.30              | 9         | 0.08                                   | 0.85 | 0.37            | –0.38              | 10        | 0.17                                   | 0.61 |
| FABER               | 0.67              | 0.42               | 17        | 0.58                                   | 0.01 | 0.58            | 0.24               | 18        | 0.23                                   | 0.34 |
| SI joint tenderness | 0.67              | 0.33               | 17        | 0.48                                   | 0.05 | 0.58            | 0.43               | 18        | 0.22                                   | 0.36 |
| ESR                 | NC                | NC                 | NC        | NC                                     | NC   | 0.36            | –0.13              | 8         | 0.71                                   | 0.05 |
| CRP                 | NC                | NC                 | NC        | NC                                     | NC   | 0.36            | 0.26               | 11        | 0.36                                   | 0.29 |
| PGA                 | 0.55              | 0.28               | 12        | 0.61                                   | 0.03 | NC              | NC                 | NC        | NC                                     | NC   |
| BASDAI              | 0.52              | 0.11               | 11        | 0.27                                   | 0.44 | 0.38            | 0.54               | 8         | 0.09                                   | 0.80 |
| BASFI               | 0.52              | –0.23              | 11        | 0.45                                   | 0.17 | 0.37            | 0.42               | 8         | 0.37                                   | 0.30 |

Change scores were calculated by subtracting final score from initial score to produce positive treatment-related change, as per convention. All 4 readings were pooled, not averaged, in calculating the SPARCC score change. SPARCC score SRM was recalculated for each clinical measure comparison by pairwise deletion of cases with missing clinical measure data. SPARCC: Spondyloarthritis Research Consortium of Canada; MRI: magnetic resonance imaging; SRM: standardized response mean; ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; BASDAI: Bath Ankylosing Spondylitis Disease Activity Index; BASFI: Bath Ankylosing Spondylitis Functional Index; PGA: physician's global assessment; SI: sacroiliac; FABER: Flexion, ABduction, and External Rotation; NC: not calculated due to insufficient observations (n < 6).

use in adults with AS or SpA and is more comprehensive than other scoring methods. This comprehensiveness is important in the evaluation of enthesitis; the anatomical concept of the “entheses organ” encompasses the functionally associated structures near the entheses itself, such as bone, insertional fibrocartilage, bursae, and synovial-covered fat pads. Beyond BME, MRI can record the intensity and 3-D extent of inflammation, as well as reflect the inflammatory changes near the entheses insertion points.

The results of our study have confirmed the reproducibility of the SPARCC scoring method for evaluating sacroiliitis in the setting of JSpA/ERA. Our results agree with available literature evaluating SPARCC's reproducibility in the adult population and further support that the SPARCC scoring method is feasible to be used in the pediatric population<sup>23,32</sup>. A reliable and objective scoring method is of great benefit for quantification of SI joint inflammation, implementation of appropriate therapies (especially anti-TNF agents), and monitoring of disease activity in children.

Most of our patients had a WB-MRI ERA protocol done, using a dedicated oblique coronal STIR sequence of SI joints adequate to detect the presence and extent of SI joint inflammation. We did not exclude any patients based on the MRI protocol used, because all patients' MRI scans had more than 6 slices through the SI joints (one of the criteria for inclusion). Our results showed high intra- and interreader reliability in scores when using a single noncontrast STIR sequence. It is our suggestion that dedicated oblique coronal sequences of SI joints should be acquired as part of the WB-MRI protocol used when the SPARCC scoring method is applied.

Our analysis did not detect any correlations or trends between the SPARCC and clinical scores, which included the PGA, BASDAI, BASFI, and CHAQ scores. This can be largely attributed to the uncommon occurrence of clinical SI involvement early in the disease course in patients with JSpA/ERA, which may often be asymptomatic and thus not recorded by physical examination. Additionally, the clinical scoring methods include but are not specific for SI involvement alone; both peripheral and axial disease activity contribute to the final scores. Peripheral involvement of joints and entheses are more frequent in JSpA/ERA patients and may have accounted for high clinical scores in patients with low SPARCC scores (little to no BME seen on the MRI). Moreover, there may be variability between clinicians in how the PGA is scored. There is also no current validated clinical score for assessing sacroiliitis alone. Future prospective studies that include standardization of the clinical scoring methods, as well as routine physical examination maneuvers of SI joint inflammation (palpation of SI joint, modified Schober test, FABER test, Gaenslen's test) even in asymptomatic patients might yield a better clinico-radiological correlation and validate the SPARCC scoring system.

The longitudinal significance of the SPARCC method could not be established in this study because of the limited number of patients who had followup MRI examinations. Importantly, just under half of the MRI studies performed on patients with JSpA/ERA demonstrated radiological evidence of sacroiliitis. Further investigation with a larger cohort and longer followup is needed to confirm the prevalence of SI involvement observed in our study, as well as to assess the

diagnostic sensitivity and specificity of the SPARCC method in the early phase of the disease.

Our study had several limitations. First, no gold standard is currently available against which the validity of MRI in detecting sacroiliitis can be compared. Synovial biopsies of the SI joint can provide pathological confirmation of the disease activity but are neither practical nor ethical to obtain in a pediatric population when less invasive forms to assess disease activity are available. Second, our present study did not record chronic radiographic disease features such as erosions, postinflammatory fatty changes, and joint space narrowing<sup>33</sup>. So it is likely that we missed chronic sacroiliitis. Current clinical assessments for sacroiliitis, including physical examination maneuvers as well as the BASDAI score, have poor sensitivity and specificity. Thus, many of the clinical outcome measures are limited in scope or driven primarily by these late changes, reflecting chronic damage and functional impairments; hence there is a mismatch in the measured construct, which in turn lowers the correlation of the 2 measurement methods.

Further, the comparison of the assessment of SPARCC and clinical outcome measures to detect response to treatment of patients' joints was limited by missing clinical data in some of the clinical measures, owing to the retrospective features of the treatment-related outcome assessment. Although the responsiveness of the SPARCC and clinical outcome scores were calculated separately for nonbiologic and biologic treatments, the treatment assignment was not random and the duration and specific drug regimen could not be stratified in this small, retrospectively examined study sample. The responsiveness of patients' joints to treatment according to clinical and MRI outcome measures in Table 3 is only comparable within each treatment group, and not between biologic versus nonbiologic treatment groups. In addition, detailed information about sports and physical activities was not assessed and documented in a standardized fashion in all patients. We could not completely rule out any association of vigorous activity and its effect on causing isolated BME. We also acknowledge that accessibility and cost of MRI may also affect its feasibility and use as a clinical tool in some centers.

Overall, our study has demonstrated high intra- and inter-reader reliability of the SPARCC scores in quantifying the level of severity of sacroiliitis in the setting of JSpA/ERA in the pediatric population. Longitudinal change in SPARCC scores consistently showed higher responsiveness to treatment-related change than most clinical outcome measures, in both the biologic and nonbiologic treatment groups. The application of this MRI scoring system in clinical practice is feasible. If the results of this pilot study are reproduced in larger series, this scoring system may serve as a reliable quantitative method to assess the degree of SI joint inflammation even in the subclinical stage and potentially to monitor disease activity and response to therapy.

## ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

## REFERENCES

1. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al; International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. *J Rheumatol* 2004;31:390-2.
2. Yilmaz M, Kendirli SG, Altintas DU, Karakoc GB, Inal A, Kilic M. Juvenile idiopathic arthritis profile in Turkish children. *Pediatr Int* 2008;50:154-8.
3. Oen K, Tucker L, Huber AM, Miettunen P, Succimarri R, Campillo S, et al. Predictors of early inactive disease in a juvenile idiopathic arthritis cohort: results of a Canadian multicenter, prospective inception cohort study. *Arthritis Rheum* 2009;61:1077-86.
4. Solau-Gervais E, Robin C, Gambert C, Troller S, Danner S, Gombert B, et al. Prevalence and distribution of juvenile idiopathic arthritis in a region of Western France. *Joint Bone Spine* 2010;77:47-9.
5. Gmucsa S, Xiao R, Brandon TG, Pagnini I, Wright TB, Beukelman T, et al. Multicenter inception cohort of enthesitis-related arthritis: variation in disease characteristics and treatment approaches. *Arthritis Res Ther* 2017;19:84.
6. Weiss PF, Klink AJ, Behrens EM, Sherry DD, Finkel TH, Feudtner C, et al. Enthesitis in an inception cohort of enthesitis-related arthritis. *Arthritis Care Res* 2011;63:1307-12.
7. Weiss PF. Diagnosis and treatment of enthesitis-related arthritis. *Adolesc Health Med Ther* 2012;3:67-74.
8. Hofer M. Spondylarthropathies in children—are they different from those in adults? *Best Pract Res Clin Rheumatol* 2006;20:315-28.
9. O'Shea FD, Boyle E, Riarh R, Tse SM, Laxer RM, Inman RD. Comparison of clinical and radiographic severity of juvenile-onset versus adult-onset ankylosing spondylitis. *Ann Rheum Dis* 2009;68:1407-12.
10. Baek HJ, Shin KC, Lee YJ, Kang SW, Lee EB, Yoo CD, et al. Juvenile onset ankylosing spondylitis (JAS) has less severe spinal disease course than adult onset ankylosing spondylitis (AAS): clinical comparison between JAS and AAS in Korea. *J Rheumatol* 2002;29:1780-5.
11. Rachlis AC, Babyn PS, Lobo-Mueller E, Benseler SM, Stimec J, Anderson M, et al. Whole body magnetic resonance imaging in juvenile spondyloarthritis: will it provide vital information compared to clinical exam alone [abstract]? *Arthritis Rheum* 2011;63:S292.
12. Burgos-Vargas R, Clark P. Axial involvement in the seronegative enthesopathy and arthropathy syndrome and its progression to ankylosing spondylitis. *J Rheumatol* 1989;16:192-7.
13. Weiss P, Beukelman T, Schanberg LE, Kimura Y, Colbert RA. Enthesitis is a significant predictor of decreased quality of life, function, and arthritis-specific pain across juvenile idiopathic arthritis (JIA) categories: preliminary analyses from the CARRAnet Registry [abstract]. *Arthritis Rheum* 2011;63:S105.
14. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, Landewé R, van der Tempel H, Mielants H, et al. Assessment of enthesitis in ankylosing spondylitis. *Ann Rheum Dis* 2003;62:127-32.
15. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 2009;68 Suppl 2:ii1-44.
16. Jans L, van Langenhove C, Van Praet L, Carron P, Elewaut D, Van Den Bosch F, et al. Diagnostic value of pelvic enthesitis on MRI of the sacroiliac joints in spondyloarthritis. *Eur Radiol* 2014;24:866-71.
17. Weiss PF. Evaluation and treatment of enthesitis-related arthritis. *Curr Med Lit Rheumatol* 2013;32:33-41.



18. Stoll ML, Bhole R, Dempsey-Robertson M, Punaro M. Spondyloarthritis in a pediatric population: risk factors for sacroiliitis. *J Rheumatol* 2010;37:2402–8.
19. Marzo-Ortega H. Interreader agreement in the assessment of magnetic resonance imaging of the sacroiliac joints in spondyloarthropathy—the 1st MISS study [abstract]. *Arthritis Rheum* 2002;46 Suppl:S428.
20. Puhakka KB, Jurik AG, Egund N, Schiottz-Christensen B, Stengaard-Pedersen K, van Overeem Hansen G, 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.
21. 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.
22. Hermann KG, Braun J, Fischer T, Reissbauer H, Bollow M. [Magnetic resonance tomography of sacroiliitis: anatomy, histological pathology, MR-morphology, and grading]. [Article in German] *Radiologe* 2004;44:217–28.
23. Landewé RB, Hermann KG, van der Heijde DM, Baraliakos X, Jurik AG, Lambert RG, et al. Scoring sacroiliac joints by magnetic resonance imaging. A multiple-reader reliability experiment. *J Rheumatol* 2005;32:2050–5.
24. Maksymowych WP, Lambert RG, Brown LS, Pagan 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.
25. Schober P. [Lendenwirbelsäule und kreuzschmerzen]. [Article in German] *Munch Med Wochenschr* 1937;84:336–8.
26. Broadhurst NA, Bond MJ. Pain provocation tests for the assessment of sacroiliac joint dysfunction. *J Spinal Disord* 1998;11:341–5.
27. Macrae IF, Wright V. Measurement of back movement. *Ann Rheum Dis* 1969;28:584–9.
28. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. Upper Saddle River: Prentice Hall; 2000.
29. Altman DG. *Practical statistics for medical research*. London: Chapman and Hall; 1990.
30. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000;53:459–68.
31. DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh* 2007;39:155–64.
32. Weber U, Maksymowych WP, Jurik AG, Pfirrmann CW, Ruffbach K, Kissling RO, et al. Validation of whole-body against conventional magnetic resonance imaging for scoring acute inflammatory lesions in the sacroiliac joints of patients with spondylarthritis. *Arthritis Rheum* 2009;61:893–9.
33. Maksymowych WP, Wichuk S, Dougados M, Jones H, Szumski A, Bukowski JF, et al. MRI evidence of structural changes in the sacroiliac joints of patients with non-radiographic axial spondyloarthritis even in the absence of MRI inflammation. *Arthritis Res Ther* 2017;19:126.