Arthritis & Rheumatism (Arthritis Care & Research)
Vol. 61, No. 7, July 15, 2009, pp 893–899
DOI 10.1002/art.24542

ORIGINAL ARTICLE

Validation of Whole-Body Against Conventional Magnetic Resonance Imaging for Scoring Acute Inflammatory Lesions in the Sacroiliac Joints of Patients With Spondylarthritis

ULRICH WEBER,[1] WALTER P. MAKSYMOWYCH,[2] ANNE G. JURIK,[3] CHRISTIAN W. A. PFIRRMANN,[1] KASPAR RUFIBACH,[4] RUDOLF O. KISSLING,[1] MUHAMMAD A. KHAN,[5] ROBERT G. W. LAMBERT,[2] AND JUERG HODLER[1]

Objective. **To compare the performance of whole-body magnetic resonance imaging (MRI) versus conventional MRI in assessing acute inflammatory lesions of the sacroiliac (SI) joints in patients with established and active spondylarthritis (SpA) using the Spondyloarthritis Research Consortium of Canada (SPARCC) MRI index. This study is validating whole-body MRI against the current MRI standard for assessing active inflammatory lesions of the SI joints in patients with SpA.**

Methods. **Thirty-two SpA patients with clinically active disease (Bath Ankylosing Spondylitis Disease Activity Index score ≥4) fulfilling the modified New York criteria were scanned by whole-body and conventional MRI of the SI joints. The MRIs were scored independently in random order by 3 readers blinded to patient identity. Active inflammatory lesions of the SI joints were recorded on a Web-based SPARCC index. Pearson's correlation coefficients were used to compare scores for whole-body and conventional MRI for each reader, whereas intraclass correlation coefficients (ICCs) were used to compare interobserver reliability.**

Results. **The Pearson's correlation coefficients between whole-body and conventional MRI per rater were 0.94, 0.87, and 0.93. The mean sum scores for conventional versus whole-body MRI were statistically significantly higher for all 3 readers, although all patients showing inflammatory lesions on conventional MRI also demonstrated them on whole-body MRI. The ICCs(2,1) were 0.69, 0.78, and 0.95 for conventional MRI, and 0.79, 0.85, and 0.96 for whole-body MRI for the 3 possible reader pairs.**

Conclusion. **Whole-body and conventional MRI scores show a strong correlation and comparable reliability for the detection of inflammatory lesions of the SI joints.**

INTRODUCTION

Spondylarthropathies (SpA) are a group of closely related chronic inflammatory rheumatic diseases, with inflammatory back pain being the leading symptom early in the disease course (1,2). Structural changes of the sacroiliac (SI) joints as defined by the modified New York classification criteria (3) may require 10 years or more before they are visible on plain radiography (4). Magnetic resonance imaging (MRI) is widely accepted as the most sensitive imaging modality to detect early inflammatory lesions of the axial skeleton (5). The 2 locations affected most frequently by inflammatory changes are the SI joints and the lower thoracic spine (6). Inflammation in SpA has been

The Whole-Body Magnetic Resonance Imaging in Spondyloarthritis project is funded by the Walter L. and Johanna Wolf Foundation, Zurich, Switzerland, and the Foundation for Scientific Research at the University of Zurich, Zurich, Switzerland. Dr. Maksymowych's work was supported by the Alberta Heritage Foundation for Medical Research.

[1]Ulrich Weber, MD, Christian W. A. Pfirrmann, MD, Rudolf O. Kissling, MD, Juerg Hodler, MD, MBA: Balgrist University Hospital, Zurich, Switzerland; [2]Walter P. Maksymowych, MD, FRCP(C), Robert G. W. Lambert, MD, FRCP(C): University of Alberta, Edmonton, Alberta, Canada; [3]Anne G. Jurik, MD: Aarhus University Hospital, Aarhus, Denmark; [4]Kaspar Rufibach, PhD: University of Zurich, Zurich, Switzerland; [5]Muhammad A. Khan, MD, MACP, FACP: Case Western Reserve University School of Medicine, MetroHealth Medical Center, Cleveland, Ohio.

Address correspondence to Ulrich Weber, MD, Department of Rheumatology, Balgrist University Hospital, Forchstrasse 340, 8008 Zurich, Switzerland. E-mail: ulrich.weber@balgrist.ch.

Submitted for publication September 21, 2008; accepted in revised form March 27, 2009.

reported as first being visible in the caudal portion of the SI joints (7–10).

Making an early diagnosis of SpA is challenging, and assessing its disease activity is mainly based on patient self-report. In recent years, MRI has gained increasing importance as a promising tool for early diagnosis of SpA and as a candidate for an objective measure of widespread inflammation (11). During the last few years, multichannel technology has been introduced into clinical MR scanners, allowing the concurrent use of several coils. This whole-body MRI method allows the scanning of the SI joints, the entire spine, the anterior chest wall, and the shoulder and pelvic girdle within 30 minutes without repositioning the patient (12). Imaging of the lower extremities is an additional option that has to be weighed against an additional examination time of 20 minutes, which may be relevant for some patients with active inflammatory back pain and peripheral enthesitis.

For the use of whole-body MRI in patients with SpA, it is important that involvement of the SI joints is validated against a current standard for imaging inflammatory changes. The goal of this study was to compare the performance of whole-body MRI versus conventional MRI in assessing acute inflammatory lesions of the SI joints in SpA patients with established and clinically active disease using the Spondyloarthritis Research Consortium of Canada (SPARCC) MRI index (13). This study design comparing 2 MRI modalities does not address the utility of whole-body MRI to diagnose SpA nor to monitor disease progression.

PATIENTS AND METHODS

Patients. Patients with active disease who were diagnosed as having SpA as defined by the modified New York criteria (3) were consecutively recruited in a single rheumatologic outpatient clinic from April 2006 to August 2007. Active disease was defined as a Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) global score of ≥4 and/or a BASDAI item 2, which measures spinal pain, score of ≥4 on a numeric rating scale ranging from 0 to 10 (14). Grading of the SI joints on plain pelvic radiographs was performed independently by 2 readers (ROK, UW) according to the radiographic modified New York criteria (3).

No limits were set with regard to the symptom duration of SpA and the age of the participants in order to include a broad spectrum of disease in this cross-sectional study. Participants were referred by practicing or hospital-based rheumatologists or by primary care physicians from many regions in Switzerland. All of the patients were also enrolled in a national prospective observational SpA cohort (Swiss Clinical Quality Management in Ankylosing Spondylitis). The following clinical and laboratory parameters were used to assess disease activity for comparisons with the MRI activity score: the complete BASDAI score; the BASDAI item 2 score; intensity and duration of lumbar morning stiffness and intensity of nocturnal pain, both assessed by a numeric rating scale ranging from 0 to 10; the Bath Ankylosing Spondylitis Functional Index (BASFI) (15); C-reactive protein (CRP) level; and erythrocyte sedimentation rate (ESR).

The protocol was approved by the Zurich Cantonal Ethics Committee, Subcommittee for Orthopedics and Rheumatology. The patients were informed in oral and written form and they provided written informed consent.

Exclusion criteria. Exclusion criteria consisted of ongoing or previous (within the last 6 months) treatment with tumor necrosis factor α inhibitors or other biologics, malignancies (including hematologic malignancies) or infections affecting the skeleton, previously performed surgery of the spine, pelvic or shoulder girdle, pregnancy, advanced spinal deformity due to SpA or other disorders precluding an adequate MRI examination, and technical contraindications to MRI such as cardiac pacemakers, neurostimulators, and similar devices.

MRI protocol. *Whole-body MRI.* Whole-body MRI was performed on a Siemens Avanto 1.5T magnet (Siemens Medical Solutions, Erlangen, Germany) with 18 independent radiofrequency channels. Various combinations of up to 6 coils plugged into the system are used depending on body position: 3 of the coils are built into the MR table (head matrix, neck matrix, spine matrix), 2 body matrix coils with 6 elements each are placed on the patient's chest and abdomen, and an additional flexible coil is placed over the hips in tall patients to increase coverage. Sagittal and coronal STIR and T1-weighted spin-echo images of the entire spine and sacrum, the anterior chest wall, and the shoulder and pelvic girdle were obtained (12). Coronal turbo STIR images were obtained with the following parameters: time to recovery (TR) 9,860 ms, time to echo (TE) 99 ms, time to inversion (TI) 130 ms, turbo factor 21, parallel acquisition technique (PAT) factor 2, and PAT mode generalized autocalibrating partially parallel acquisition (GRAPPA). Coronal T1-weighted spin-echo sequence images were obtained with the following parameters: TR 571 ms, TE 12 ms, PAT factor 2, and PAT mode GRAPPA. Two imaging steps were used with a field of view (FOV) of 450 × 450 mm and an imaging matrix of 660 × 384 pixels per step, 5 mm section thickness, interslice gap 1.0 mm, and 32 sections, resulting in an FOV of 780 × 450 mm. The sum of the acquisition times for all sequences, including 2 localizers, was 21 minutes 53 seconds. The total acquisition time, including patient positioning, was 30 minutes.

Conventional MRI. For conventional MRI, angled coronal turbo STIR and turbo T1-weighted spin-echo images and angled transverse turbo STIR images of both SI joints were obtained. Imaging parameters for both STIR sequences were TR 4,930 ms, TE 69 ms, TI 150 ms, turbo factor 9, PAT factor 2, and PAT mode GRAPPA, with a matrix size of 256 × 256 pixels. The parameters of the coronal T1 sequence were TR 450 ms, TE 12 ms, turbo factor 3, PAT factor 2, PAT mode 2, and GRAPPA 2, with a matrix size of 512 × 256 pixels. In all 3 sequences, the FOV was 280 × 280 mm and slice thickness was 4 mm, with an interslice gap of 0.4 mm. Nineteen sections were obtained per sequence. The sum of sequence acquisition

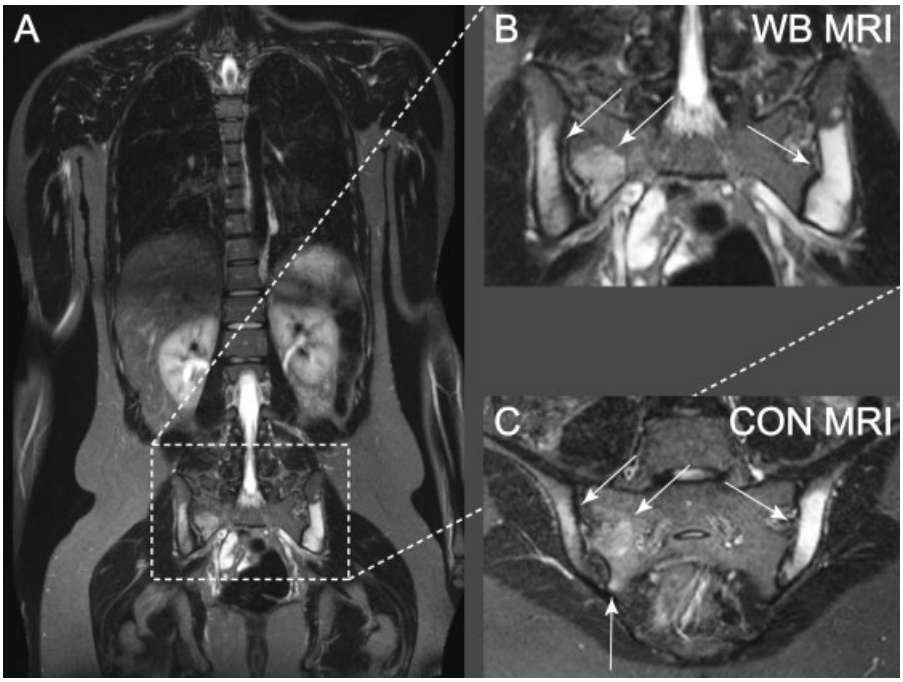

Figure 1. Magnetic resonance image (MRI) of a female patient age 26 years: HLA–B27 positive, symptoms duration 4 years, erythrocyte sedimentation rate 16 mm/hour, Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) global score 5.4, and BASDAI item 2 score 6. **A**, Coronal whole-body (WB) MRI, STIR sequence; active inflammatory lesions show in both sacroiliac (SI) joints (**insert**). **B**, Coronal WB MRI, detail SI joints, STIR sequence: bilateral sacroiliitis (**arrows**) on the right side, involving both the iliac and sacral part of the joint. **C**, Conventional (CON) MRI of the SI joints, STIR sequence: corresponding bilateral sacroiliitis (**arrows**) in a dedicated oblique (semicoronal) SI joint slice.

times was 7 minutes 21 seconds. With the time required by the system to start the next sequence and assuming a normally mobile patient, the total acquisition time for conventional MRI of the SI joints, including patient positioning, is ~15 minutes.

The difference in slice thickness between whole-body and conventional MRI is explained by the fact that for whole-body MRI, the entire anteroposterior diameter of the body from the anterior chest wall and symphysis to the posterior elements of the spine and the SI joints has to be covered. For conventional MRI, only the anteroposterior diameter of the SI joints needs to be included. This is reached by a combination of increased slice thickness and the number of slices in whole-body MRI.

Analysis of MRIs. Quality checks of each MR scan were performed, and patients with complete sets of whole-body and conventional MRIs were enrolled in the analysis of the images (Figure 1).

The whole-body and conventional MRIs were read and scored independently by 3 readers (AGJ, WPM, JH) who were not involved in the clinical examination of the patients and who were blinded to patient identity and clinical parameters. One reader (AGJ, radiologist) is an experienced Outcome Measures in Rheumatology Clinical Trials (OMERACT) reader and one reader (WPM, rheumatologist) is a coauthor of the SPARCC MRI index (13); the third reader (JH, radiologist) participated in a training and calibration exercise by videoteleconference sessions based on reference cases provided by the Canada-Denmark MRI in SpA Working Group. Each reader also reviewed an online training module describing the approach to the scoring of acute inflammation in the SI joints using the SPARRC SI joints scoring method (16).

The films were reviewed in random order (regarding the sequence of patients as well as the sequence of whole-body and conventional MRI using a Web-based random number generator) on electronic work stations in the institution of each reader. Scores were recorded electronically on a separate screen using a masked reading program.

Active inflammatory lesions on STIR sequences of the SI joints on whole-body and conventional MRI films were scored using a Web-based scoring program developed specifically for the SPARCC MRI index (online at www.sparccmri.filipow.ca). The SPARCC method assesses 6 consecutive coronal slices of the SI joints, which permits a 3-dimensional assessment of the extent of the lesion. Both SI joints are scored in 4 quadrants (superior and inferior iliac and sacral quadrants). With additional scores assigned to lesions exhibiting depth and/or intensity, the total SI joint score ranges from 0 to 72. In this study, however, we did not assign an additional score for intensity because this may partly depend on the technical specifications of the MRI device. In addition to selecting the 6 most severely affected slices, we scored all slices demonstrating an increased STIR signal.

Statistical analysis. To assess the association between whole-body and conventional MRI scores within a single rater, scatterplots of these respective measurements were inspected and Pearson's correlation coefficients were computed. Confidence intervals (CIs) for Pearson's correlation coefficients were generated using Fisher's z-transformation. Mean sum scores were compared using the Mann-Whitney test. To check for systematic relationships between the differences of the conventional and whole-body measurements and their means, increasing variability, and potentially necessary transformation of reader scores within each reader, Bland-Altman plots with corresponding limits of agreement were inspected (17). The average score of each rater for the 2 MRI methods on the x-axis was plotted against the score difference (conventional − whole-body) on the y-axis, including 95% limits of agreement. No need for any transformation of score measurements was detected. In all of these analyses, no correction for multiple testing was performed and the alpha significance level was set to 0.05. The intrareader correlation was defined as moderate, good, very good, and excellent by the values >0.5, >0.6, >0.8, and >0.9, respectively.

Within each method, intraclass correlation coefficients (ICCs) were computed based on a 2-way random-effects model with single measurements to compare interobserver reliability (18). The rationale for using this model is that readers are considered to be a random sample from the population of all readers; this ICC variant is commonly abbreviated as ICC(2,1). Consequently, the results are representative of a larger population of raters. The ICC variant ICC(3,1) was additionally calculated; this approach considers the raters to be fixed (because 2 readers had a calibration exercise for scoring prior to the study readout) and therefore not representative of a larger population. When considering an "ICC(a,b)," the "a" refers to 1 of 3 possible models used (19). If "a" equals 2, each patient is rated by each reader and readers are randomly selected and therefore representative of a larger population of readers. If "a" equals 3, raters are considered to be fixed and the only raters of interest, and therefore not representative of a larger population. The "b" in "ICC(a,b)" can take any positive integer value and represent the form of the ICC. The form reflects whether the reliability is to be calculated on a single measurement or by taking the average of "b" measurements taken by different raters.

ICCs were computed for each reader pair and for all 3 readers jointly within each site, MRI method (conventional versus whole-body MRI), and number of slices categories (6 most affected versus all affected sections). ICCs >0.5, >0.6, >0.8, and >0.9 were regarded to represent a moderate, good, very good, and excellent reproducibility, respectively.

Under the additional assumption that we are looking at a random sample from a population of trained readers, we also provide CIs for the ICC(2,1) (20). Since we were only interested in its lower limit, these CIs were computed using an alpha significance level of 0.1. Given such a CI, we can claim that the population ICC is not smaller than the lower bound of the CI, with a probability of ≥95%. According to Rousson et al (20), an ICC can be considered good once the lower limit of the 90% confidence interval (90% CI) is ≥0.75.

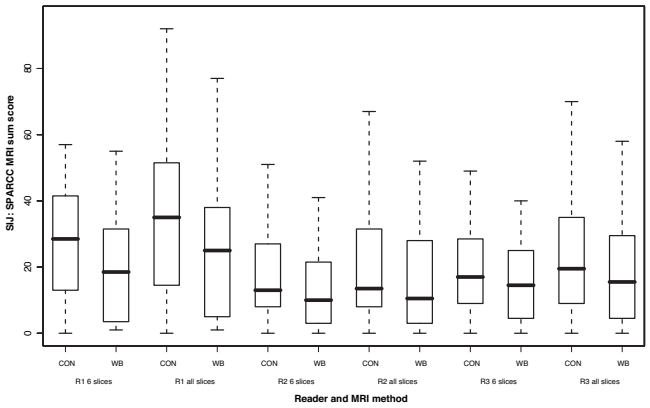

Figure 2. Spondyloarthritis Research Consortium of Canada (SPARCC) magnetic resonance imaging (MRI) sacroiliac joint (SIJ) scores for the 3 readers, for the 6 most affected SIJ slices and for all of the affected SIJ slices, and for both conventional (CON) and whole-body (WB) MRI methods. R1 = reader 1; R2 = reader 2; R3 = reader 3.

RESULTS

Patient characteristics. Thirty-two patients (27 men) with SpA fulfilling the modified New York criteria were recruited; 30 patients had primary ankylosing spondylitis and 2 had associated Crohn's disease. The median age was 35.5 years (range 17.3–65.5) and the median symptoms duration was 10.5 years (range 1–37). Sixteen patients (50%) had symptoms duration of ≤10 years. Twenty-seven patients were found to be HLA–B27 positive; in 1 patient this genetic marker had not been determined.

The median disease activity as measured by the patient self-reported BASDAI global score was 4.3 (range 2–7.2), and the median BASDAI item 2 score was 6.5 (range 4–10). The median ESR was 15 mm/hour (range 2–60) and the median CRP level was 6.5 mg/liter (0.5–46). The BASFI ranged from 0 to 7.7, with a median value of 3.5. The median Bath Ankylosing Spondylitis Metrology Index score (21) was 1.0 (range 0–7).

Intrareader correlation for whole-body and conventional MRI scores of the SI joints. The SPARCC MRI index SI joint scores for the 3 readers and for both MRI methods are shown in Figure 2. The majority of observations lay within the limits of agreement in the Bland-Altman plots. The Pearson's correlation coefficients between whole-body and conventional MRI per rater for the 3 readers were 0.90, 0.87, and 0.90 for the 6 most affected SI joint slices and 0.94, 0.87, and 0.93 for all SI joint slices with inflammatory lesions (Table 1).

The mean sum scores of all 3 readers were statistically significantly higher for conventional MRI compared with whole-body MRI, both for the 6 most affected sections and for all SI joint slices displaying inflammatory lesions (Table 1). On an individual patient level, all of the patients with inflammatory lesions of the SI joints in conventional

Table 1. Intrareader correlation for whole-body and conventional MRI scores of the SI joints*

	Reader 1	Reader 2	Reader 3
Pearson's correlation coefficient			
All slices†	0.94	0.87	0.93
6 slices‡	0.90	0.87	0.90
Mean sum scores			
All slices†			
Conventional MRI	35.4	21	23.5
Whole-body MRI	25	16.2	18.5
Mann-Whitney test *P*	< 0.0001	0.0005	0.0004
6 slices‡			
Conventional MRI	27.4	18.1	19.4
Whole-body MRI	19.7	14.4	16.1
Mann-Whitney test *P*	< 0.0001	0.0018	0.0020

* MRI = magnetic resonance imaging; SI = sacroiliac.
† All SI joint slices with inflammatory lesions.
‡ The 6 slices most affected by inflammatory changes.

MRI also showed inflammatory changes in whole-body MRI and vice versa.

Interobserver reliability for whole-body and conventional MRI scores of the SI joints. The interobserver reliability as expressed by the ICC(3,1), which considers the observers as a fixed sample and therefore not representative of a larger population of readers, ranged for the 3 reader pairs from 0.80 to 0.96 for conventional MRI and from 0.87 to 0.97 for whole-body MRI (Table 2). The ICC(2,1), which considers the readers as a random sample and representative of a larger population of observers, resulted in ranges for the 3 reader pairs from 0.69 to 0.96 for conventional MRI and from 0.79 to 0.96 for whole-body MRI. The interobserver reliability for the reader pair 2/3, who participated in a calibration exercise prior to the study readout, was higher for both MRI techniques (ranges for both ICC variants were 0.95–0.96 for conventional MRI and 0.96–0.97 for whole-body MRI).

The 90% CIs for ICC(2,1) were derived by making the extra assumption of a population of trained readers. The lower limits of the 90% CI for all raters jointly and for the 6 most affected SI joint slices were 0.76 for whole-body MRI and 0.66 for conventional MRI, and for all affected SI joint sections were 0.74 for whole-body MRI and 0.68 for conventional MRI. The lower bounds of the 90% CI for the reader pair 2/3 trained by a calibration exercise for the 6 most affected SI joint slices were 0.88 for whole-body MRI and 0.88 for conventional MRI, and for all affected SI joint sections were 0.90 for whole-body MRI and 0.86 for conventional MRI.

Correlation of whole-body and conventional MRI scores of the SI joints with clinical variables. All but 1 of 96 Pearson's correlations between mean SI joint MRI scores from all 3 readers (whole-body or conventional, 6 most affected or all affected slices) and BASDAI global score, BASDAI item 2 score (axial pain), BASFI, ESR and CRP level, and nocturnal pain or morning stiffness (intensity and duration) were not significant (data not shown). Given the large number of statistical tests, we consider the sole significant correlation as a false-positive.

DISCUSSION

This study demonstrated that inflammatory SI joint lesions in patients with active SpA are highly correlated between conventional MRI and the recently introduced whole-body MRI technique. The patient sample in this study was representative of daily routine in an outpatient department of rheumatology. The interreader agreement was good to excellent, depending on the statistical assumptions concerning the characteristics of the reader team and on precalibration activities prior to the readout.

A whole-body MRI examination to assess the entire axial skeleton takes only ~10 minutes longer to perform than conventional MRI of a limited region. The multichannel technology combined with a multicoil system and dedicated software to fuse the regional MRI examination results in an equal spatial resolution of whole-body compared with conventional MRI. With its comprehensive examination, whole-body MRI was first introduced in the

Table 2. Interobserver reliability for whole-body and conventional MRI scores of the SI joints*

	RP 1/2	RP 1/3	RP 2/3†	All readers
Conventional MRI				
All slices, ICC(2,1)/ICC(3,1)	0.69/0.82	0.78/0.88	0.95/0.96	0.78/0.88
All slices, 90% CI LB for ICC(2,1)	0.54	0.66	0.86	0.68
6 slices, ICC(2,1)/ICC(3,1)	0.69/0.80	0.77/0.86	0.96/0.96	0.79/0.87
6 slices, 90% CI LB for ICC(2,1)	0.50	0.61	0.88	0.66
Whole-body MRI				
All slices, ICC(2,1)/ICC(3,1)	0.79/0.87	0.85/0.89	0.96/0.97	0.85/0.91
All slices, 90% CI LB for ICC(2,1)	0.64	0.68	0.90	0.74
6 slices, ICC(2,1)/ICC(3,1)	0.84/0.89	0.87/0.89	0.96/0.96	0.88/0.91
6 slices, 90% CI LB for ICC(2,1)	0.69	0.68	0.88	0.76

* All slices are all of the SI joint slices with inflammatory lesions; 6 slices are the 6 slices most affected by inflammatory changes. MRI = magnetic resonance imaging; SI = sacroiliac; RP = reader pair; ICC = intraclass correlation coefficient; ICC(2,1) = considering the raters as a random sample; ICC(3,1) = considering the raters as a fixed sample; CI LB = confidence interval, lower 90% bound.
† Participated in a prereading calibration exercise by a videoteleconference session.

specialities of oncology (22) and angiology (23). However, it may also prove useful in systemic musculoskeletal disorders such as SpA. We therefore performed this validation study assessing inflammatory lesions of the SI joints by whole-body and conventional MRI.

The intrareader correlation between whole-body and conventional MRI scores measured by Pearson's correlation coefficients was very good, with values ranging from 0.87 to 0.94 for all SI joint slices with inflammatory lesions and from 0.87 to 0.90 for the 6 most affected SI joint sections. The significantly higher scores for conventional MRI for all readers are probably due to the different SI joint slice orientations with strictly coronal sections in whole-body MRI compared with oblique slices in conventional MRI that depict the synovial portion of the joint. We hypothesize that this difference in slice orientation may have resulted in a larger area of the SI joints scanned by conventional MRI. Another reason that may partially explain this observation is the difference in SI joint slice thickness between the 2 MRI techniques (5 mm in whole-body MRI versus 4 mm in conventional MRI). However, all patients with inflammatory lesions of the SI joints in conventional MRI also showed them on whole-body MRI and vice versa.

Interobserver reliability is often expressed by ICCs. At least 6 different approaches exist to compute ICCs, depending on whether readers are considered a random or fixed sample (meaning representative of a larger population of readers or not), whether each case is assessed by each reader or not, and whether ratings are individual reader values or averages over several readers. Our 3 readers consisted of an experienced OMERACT radiologist, a rheumatologist and cofounder of the SPARCC MRI scoring index, and an experienced musculoskeletal radiologist involved in the adaptation of the whole-body MRI technique for SpA. The ICC(2,1) variant, which considers the readers to be representative of a larger population of observers, resulted in a broader range from 0.69 to 0.96 for the 3 reader pairs, both for all affected and the 6 most affected SI joint sections; 5 of 16 possible combinations of 3 reader pairs and all readers jointly, and 2 MRI methods were seen with a 90% CI lower limit of $\geq$0.75. The analysis of the same data set by the ICC(3,1) variant, which considers readers to not be representative of a larger population of observers, yielded a smaller range for interobserver reliability of 0.80–0.97. This suggests that our reader team may not be representative of a larger population of observers.

Prior to the readout, the reader pair 2/3 had the opportunity of a training and calibration exercise by videoteleconference sessions using reference images and whole-body MRI examinations of patients not enrolled in this study. This resulted in a very high agreement between these 2 readers of 0.95–0.96 using the ICC(2,1) and 0.96–0.97 using the ICC(3,1), and the 90% CI lower limit of $\geq$0.75 criterion was fulfilled (range 0.86–0.90) for all 4 possible combinations (conventional and whole-body MRI, 6 most affected and all affected SI joint slices). Based on this observation, we recommend a prior calibration session for all participating readers before starting a study readout with multiple observers. A multireader exercise comparing 5 different MRI indices for scoring SI joints showed a substantial interreader variability among 7 different not specifically trained observers with ICCs(2,1) per reader pair, ranging from 0.13 to 0.86 for status scores (24).

In line with a growing number of previous studies assessing disease activity by inflammatory lesions in MRI (25–27), we were not able to detect a correlation of SI joint scores by either MRI technique with several clinical and laboratory parameters for disease activity (BASDAI global score, BASDAI item 2 score, intensity and duration of morning stiffness, nocturnal pain, BASFI, ESR, and CRP). We speculate that the assessment of inflammatory SI joint lesions in MRI reflects other aspects of disease activity than the ones expressed by these clinical and laboratory parameters.

Whole-body and conventional MRI showed a very good correlation for the detection of inflammatory lesions in the SI joints in patients with established and active SpA. However, a recommendation to use whole-body instead of conventional MRI in daily routine seems to be premature. Further studies need to address the comparative performance for assessing inflammatory lesions in the spine and the clinical relevance of the additional information on inflammation in the anterior chest wall and the hip and shoulder girdles. With its more comprehensive assessment of inflammation compared with conventional MRI, whole-body MRI may serve as an objective and quantitative measure of inflammatory lesions in the entire axial skeleton in SpA. These properties of whole-body MRI, which requires an additional examination time of only ~10 minutes compared with conventional MRI of the SI joints or a selected spinal region, may also prove useful to define the sensitivity and specificity of inflammatory lesions in different regions of the axial skeleton. Such a validation of inflammatory lesions would provide the basis to use MRI for diagnostic classification in early SpA. Complementary to analyzing status scores, the validation process of whole-body MRI needs to also compare responsiveness and reliability of change scores.

In conclusion, all patients with SpA showing inflammatory lesions of the SI joints on conventional MRI also demonstrated them on whole-body MRI, as measured by the SPARCC Index, and the 2 MRI methods showed a very good correlation. The interreader reliability was good to excellent depending on the statistical assumptions concerning the characteristics of the reader team and depending on precalibration prior to the readout.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Weber had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Weber, Maksymowych, Jurik, Pfirrmann, Kissling, Khan, Lambert, Hodler.

Acquisition of data. Weber, Maksymowych, Jurik, Pfirrmann, Kissling, Hodler.

Analysis and interpretation of data. Weber, Maksymowych, Jurik, Pfirrmann, Rufibach, Kissling, Khan, Lambert, Hodler.

ACKNOWLEDGMENTS

We thank the patients for their participation; Désirée van der Heijde, MD, PhD, Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands, for her advice concerning the study design; Tracey Clare, Clinical Research Manager, and Paul Filipow, Data Manager, Department of Radiology, University of Alberta, Edmonton, Canada for coordinating the Web-based SPARRC scoring index and Christian Streng, Balgrist University Hospital, Zurich, Switzerland for his technical assistance with Figure 1. We also thank the following Swiss rheumatologists, internists, and primary care physicians for referral of their patients: D. Amgwerd, Spreitenbach; G. Bickel, Rapperswil; C. Boetschi, Romanshorn; C. Brunner, T. Gerber, F. Haefelin, G. Hajnos, J. Imholz, I. Kramers, P. Sutter, Zurich; S. Buergin, A. Schmidt, B. Weiss, Basel; P. De Vecchi, St. Moritz; D. Galovic, Pfaeffikon; M. Giger, Menzingen; C. Gut, Reinach; C. Jeanneret, Schwerzenbach; M. Klopfstein, Biel; A. Meniconi, Schwyz; S. Pfister, Buelach; J. Sturzenegger, Kreuzlingen; F. Tapernoux, Rueti; R. Wuethrich, Brugg.

REFERENCES

1. Khan MA. Update on spondyloarthropathies. Ann Intern Med 2002;136:896–907.
2. Braun J, Sieper J. Ankylosing spondylitis. Lancet 2007;369: 1379–90.
3. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. Arthritis Rheum 1984; 27:361–8.
4. Mau W, Zeidler H, Mau R, Majewski A, Freyschmidt J, Stangel W, et al. Clinical features and prognosis of patients with possible ankylosing spondylitis: results of a 10-year followup. J Rheumatol 1988;15:1109–14.
5. Weber U, Maksymowych WP. How does imaging help the clinician in the evaluation and management of spondyloarthritis? Skeletal Radiol 2008;37:487–90.
6. Baraliakos X, Landewe R, Hermann KG, Listing J, Golder W, Brandt J, et al. Inflammation in ankylosing spondylitis: a systematic description of the extent and frequency of acute spinal changes using magnetic resonance imaging. Ann Rheum Dis 2005;64:730–4.
7. Muche B, Bollow M, Francois RJ, Sieper J, Hamm B, Braun J. Anatomic structures involved in early- and late-stage sacroiliitis in spondylarthritis: a detailed analysis by contrast-enhanced magnetic resonance imaging. Arthritis Rheum 2003; 48:1374–84.
8. Puhakka KB, Jurik AG, Schiottz-Christensen B, Hansen GV, Egund N, Christiansen JV, et al. MRI abnormalities of sacroiliac joints in early spondylarthropathy: a 1-year follow-up study. Scand J Rheumatol 2004;33:332–8.
9. Puhakka KB, Melsen F, Jurik AG, Boel LW, Vesterby A, Egund N. MR imaging of the normal sacroiliac joint with correlation to histology. Skeletal Radiol 2004;33:15–28.
10. Weber U, Pfirrmann CW, Kissling RO, MacKenzie CR, Khan MA. Early spondyloarthritis in an HLA-B27-positive monozygotic twin pair: a highly concordant onset, sites of involvement, and disease course. J Rheumatol 2008;35:1464–7.
11. Weber U, Kissling RO, Hodler J. Advances in musculoskeletal imaging and their clinical utility in the early diagnosis of spondyloarthritis. Curr Rheumatol Rep 2007;9:353–60.
12. Weber U, Pfirrmann CW, Kissling RO, Hodler J, Zanetti M. Whole body MR imaging in ankylosing spondylitis: a descriptive pilot study in patients with suspected early and active confirmed ankylosing spondylitis. BMC Musculoskelet Disord 2007;8:20.
13. 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.
14. Garret S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
15. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21: 2281–5.
16. Chiowchanwisawakit P, Pedersen SJ, Lambert RG, Dhillon SS, Filipow P, Maksymowych WP. Validation of an online training module for scoring inflammation in the sacroiliac joints by the SPARCC MRI method [abstract]. Ann Rheum Dis 2008;67 Suppl II:512.
17. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.
18. Rousson V, Gasser T, Seifert B. Assessing intrarater, interrater and test-retest reliability of continuous measurements. Stat Med 2002;21:3431–46.
19. Shrout PE, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. Psychol Bull 1979;86:420–8.
20. Rousson V, Gasser T, Seifert B. Confidence intervals for intraclass correlation in interrater reliability. Scand J Statist 2003;30:617–24.
21. Jenkinson TR, Mallorie PA, Whitelock HC, Kennedy LG, Garrett SL, Calin A. Defining spinal mobility in ankylosing spondylitis (AS): the Bath AS Metrology Index. J Rheumatol 1994; 21:1694–8.
22. Schaefer JF, Schlemmer HP. Total-body MR-imaging in oncology. Eur Radiol 2006;16:2000–15.
23. Nael K, Ruehm SG, Michaely HJ, Saleh R, Lee M, Laub G, et al. Multistation whole-body high-spatial-resolution MR angiography using a 32-channel MR system. Am J Roentgenol 2007; 188:529–39.
24. Landewe 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.
25. Puhakka KB, Jurik AG, Schiottz-Christensen B, Hansen GV, Egund N, Christiansen JV, et al. Magnetic resonance imaging of sacroiliitis in early seronegative spondyloarthropathy: abnormalities correlated to clinical and laboratory findings. Rheumatology (Oxford) 2004;43:234–7.
26. Maksymowych WP, Dhillon SS, Park R, Salonen D, Inman RD, Lambert RG. Validation of the Spondyloarthritis Research Consortium of Canada magnetic resonance imaging spinal inflammation index: is it necessary to score the entire spine? Arthritis Rheum 2007;57:501–7.
27. Lambert RG, Salonen D, Rahman P, Inman RD, Wong RL, Einstein SG, et al. Adalimumab significantly reduces both spinal and sacroiliac joint inflammation in patients with ankylosing spondylitis. Arthritis Rheum 2007;56:4005–14.