Diagnostic Utility of Magnetic Resonance Imaging and Radiography in Juvenile Spondyloarthritis: Evaluation of the Sacroiliac Joints in Controls and Affected Subjects

Jacob L. Jaremko, Lei Liu, Naomi J. Winn, Janet E. Ellsworth, and Robert G.W. Lambert

ABSTRACT. Objective. To compare the utility of radiography and magnetic resonance imaging (MRI) for the diagnosis of juvenile-onset spondyloarthritis in pediatric patients presenting with low back and/or sacroiliac (SI) pain of potentially inflammatory etiology.

Methods. Radiographs and MRI studies of the SI joints in 26 patients with juvenile spondyloarthritis (JSpA) and 35 controls were assessed independently by 2 radiologists, with discrepancies arbitrated by a third. Radiographs and MRI were blinded and read in separate batches in random order.

Results. Erosion was common and was the most useful diagnostic feature on radiography [positive likelihood ratio (LR) = 3.5] and was especially diagnostic of SpA on MRI (LR = 6.7). Subchondral sclerosis was common but was the least specific feature for both modalities. Joint space narrowing had some utility on radiography (LR = 2.0) and MRI (LR = 2.7) but was uncommon and had poor reader reliability. Bone marrow edema (LR = 3.1) and subarticular fat infiltration (LR = 4.5), detectable only on MRI, were both useful features. Global diagnostic impression of MRI (LR = 9.4) had very high utility for the diagnosis of JSpA, exceeding radiography (LR = 4.4) because of superior specificity. In addition, global diagnosis of SpA is much more reliably made on MRI (κ = 0.80) compared to radiography (κ = 0.30).

Conclusion. Specificity and reliability of MRI of the SI joints are superior to radiography for the diagnosis of juvenile-onset SpA and, where available, MRI should replace radiography as the first line of investigation. (First Release March 15 2014; J Rheumatol 2014;41:963–70; doi:10.3899/jrheum.131064)

Key Indexing Terms: SPONDYLARTHROPATHIES ANKYLOSING SPONDYLITIS RADIOGRAPHY MAGNETIC RESONANCE IMAGING PEDIATRICS

From the Department of Radiology and Diagnostic Imaging, and the Department of Pediatrics, W.C. Mackenzie Health Sciences Centre, University of Alberta, Edmonton, Alberta, Canada.

J.L. Jaremko, MD, PhD, FRCPC; L. Liu, MD; N.J. Winn, MBBS.

BnedSci, Department of Radiology and Diagnostic Imaging, W.C. Mackenzie Health Sciences Centre, University of Alberta; J.E. Ellsworth, MD, FRCP, Department of Pediatrics, University of Alberta; R.G.W. Lambert, MB, BCH, FRCP, Department of Radiology and Diagnostic Imaging, W.C. Mackenzie Health Sciences Centre, University of Alberta. Address correspondence to Dr. J.L. Jaremko, Department of Radiology and Diagnostic Imaging, 2A2.41 WC Mackenzie Health Sciences Centre, University of Alberta, 8440-112 Street, Edmonton, Alberta T6G 2B7, Canada. E-mail: jjaremko@ualberta.ca

Accepted for publication December 24, 2013.

Up to one-fifth of cases of spondyloarthritis (SpA) have onset before age 16, termed juvenile spondyloarthritis (JSpA)¹,². Compared to adults, JSpA more often presents with peripheral arthritis and enthesitis, while symptoms involving the spine and sacroiliac (SI) joints often occur later³,⁴,⁵. Given this presentation, clinical diagnosis of JSpA can be difficult, and the imaging diagnosis may be more critical. Historically, SI joints have been imaged by radiography; however, the structural damage that radiographs detect is often not present or below detection threshold in early disease⁶. Increasingly, magnetic resonance imaging (MRI) of the SI joints is being done for earlier diagnosis.

MRI assessment of SpA is well supported by the literature in adults, with consensus definitions of lesions indicating pathology on MRI now incorporated into diagnostic criteria for SpA⁷,⁸. In contrast, there is less published evidence regarding MRI assessment of sacroiliitis in children. Previous studies showed MRI to be more sensitive for sacroilitis than radiography in children⁹, using gadolinium contrast-enhanced protocols that are not routinely performed in most centers. More recent studies of imaging findings lacked control subjects and so were unable to determine diagnostic utility¹⁰,¹¹. These studies also showed essentially no radiographic abnormality in patients with established JSpA, a finding contrary to our experience. Because the available evidence regarding the normal range of appearances of the SI joints in a juvenile population is limited, the significance of findings on MRI of SI joints is unclear in some cases. So we assessed the utility of MRI and radiography in the diagnosis of JSpA by comparing imaging
findings at the initial MRI performed for assessment of symptoms in patients who were ultimately diagnosed with JSpA compared to control subjects.

MATERIALS AND METHODS

Our study is an ethics-approved retrospective cohort analysis done at a tertiary referral pediatric hospital.

Population. Regional imaging archives between March 2005 and May 2012 were searched for patients who had (1) MRI of the SI joints performed at age ≤ 18 years; (2) radiographs depicting the SI joints performed at age ≤ 18 and within 2 years of the MRI study; and (3) a clearly established final clinical diagnosis. Patients were excluded if the SI joints could not be fully visually assessed because of missing or poor-quality images or pathological disruption such as obliteration by tumor. The patient cohort included those who had MRI of the SI joints for symptoms of back pain or SI joint pain, whether or not this pain was known to meet Assessment of SpondyloArthritis criteria for inflammatory back pain. They also (1) had onset at age ≤ 16 years of disease that met International League of Associations for Rheumatology criteria for juvenile idiopathic arthritis (JIA) of the enthesitis-related, psoriatic, or undifferentiated subtypes (because these subtypes include sacroiliitis as a possible classification criterion); or (2) had inflammatory bowel disease (IBD) associated with low back and/or SI joint pain; or (3) had clinical findings of inflammatory low back/SI joint pain in combination with clinical and imaging evidence of sacroiliitis that indicated presence of JSpA in the expert opinion of a rheumatologist. The diagnosis of the original community rheumatologist was subsequently verified on chart review by a pediatric rheumatologist using the combination of patient history, clinical examination, laboratory data, imaging findings, and longterm clinical followup. This cohort ultimately included 26 patients: 12 with SpA/ankylosing spondylitis, 5 with JIA (enthesitis-related arthritis subtype), 4 with reactive arthritis, 4 with IBD associated arthritis/SpA, and 1 with undifferentiated JIA. Subjects who had a clear clinical diagnosis other than SpA were placed into the control group. This comprised 35 patients, with 13 cases of trauma, 10 of mechanical back pain, 3 of JIA of oligoarticular or polyarticular subtypes (involving only peripheral joints and specifically excluding sacroiliitis), and miscellaneous etiologies including fibromyalgia, lumbar pars interarticularis defect, spondylolisthesis, and cystic fibrosis arthropathy.

Imaging and assessment. Images were retrieved from the regional imaging database, anonymized, and randomized for blinded assessment using ClearCanvas Workstation. Eligible radiographs were either anteroposterior views of the pelvis or dedicated SI joint projections. All available projections were assessed. MRI scans were all performed locally on 1.5T magnets and included at a minimum 3 primary sequences recommended for SI joint imaging, including coronal T1 and short-tau inversion recovery (STIR), and transverse STIR images oriented parallel and perpendicular to the SI joints respectively, i.e., oblique with respect to the MRI scanner axes. We did not routinely obtain sequences with gadolinium enhancement or dedicated to cartilage mapping.

Two primary readers independently assessed all available images. Discordant results were arbitrated by a third independent reader who was also blinded to all information. Primary readers were a musculoskeletal radiology fellow near completion of training, and a staff radiologist with dual pediatric and musculoskeletal radiology fellowships. The arbitrator was a senior musculoskeletal radiologist with 30 years of experience of SpA imaging.

Observers read each modality separately, blinded to patient identifiers, clinical history, and the other modality. Imaging studies were randomized and radiographs were read first. After at least 1 week, MRI were read. For both modalities, the presence of sclerosis, erosion, and articular narrowing was assessed. On MRI, bone marrow edema (BME) and subarticular fat infiltration were also evaluated.

Definitions for disease findings were taken from the adult literature with minor adaptations for a juvenile population. The definitions were standardized and developed through an iterative training exercise using images from patients excluded from the study (most commonly because of only 1 imaging modality being available). The following is a brief summary of the definitions, with additional detail for the definitions of sclerosis and erosion presented in Figures 1 and 2. On radiography, sclerosis was defined as an increase in subchondral bone density greater than at adjacent joints or disk endplates; erosion as an irregularity of the osteochondral interface; and narrowing as a decrease in the expected joint width. On MRI, sclerosis was defined on all sequences as a much lower subarticular signal than normal and/or a wider than normal area of very low signal; erosion as irregularity in the osteochondral interface involving both contour and signal\(^1\)\(^2\); and narrowing as decrease in expected joint width. Features visible only on MRI were defined as published\(^3\)\(^4\) and included BME — an ill-defined increase in marrow STIR signal without alternative explanation; and fatty infiltration — focal subarticular T1 signal greater than normal for age, usually accompanied by decreased STIR signal. MRI definitions for sclerosis and erosion were also visually defined for additional clarity (Figure 1C and 2C). All features were scored in binary fashion (present or absent).

For radiographs and MRI, readers also separately recorded a binary (yes/no) global impression in response to the statement “This study confirms the presence of spondylololithitis,” and their level of confidence in this response on a scale from ~5 (definitely not SpA) to +5 (definitely SpA), with 0 representing equipoise.

Statistical analysis. The frequency of patients with JSpA and controls who had specific radiographic and MRI abnormalities and the global impression of diagnosis were analyzed descriptively according to the concordant/arbitrated result. Decision matrix analysis was carried out to generate sensitivity (SN), specificity (SP), and likelihood ratios (LR) for each disease finding and global impression as individual predictors of clinical diagnosis. Given the sample sizes available, we used Fisher’s exact test to assess for statistically significant differences (at p = 0.05) in the distribution of disease features between disease and control groups. Interreader variability was established between the 2 primary readers using Cohen’s k test and percent agreement. Lastly, confidence scores of individual readers were compared between the modalities using the Student’s t-test, for both positive and negative diagnoses of JSpA.

RESULTS

Demographically, the JSpA cohort was 85% male, while the control cohort was 46% male. Other characteristics were similar between JSpA and control groups: mean ± SD age at time of study was 14.9 ± 2.3 and 15.1 ± 2.4 years for radiography and MRI in patients with JSpA, respectively, versus 15.0 ± 2.1 and 15.1 ± 2.3 years in controls; time interval between radiography and MRI studies was 5 ± 4 and 4 ± 4 months in the JSpA and control groups, respectively. In the JSpA group, age at disease onset was 12.9 ± 2.6 years. Among patients with JSpA, 15/20 (75%) with available HLA-B27 test results were positive for that allele. Inflammatory markers were considered relevant if tested within 6 months of MRI. For disease patients with available test results, the erythrocyte sedimentation rate (ESR) was abnormally elevated (> 15 mm/h) in 14/24 (58%), and C-reactive protein levels were elevated (> 8.0 mg/l) in 13/23 (57%).

Global impression for the diagnosis of SpA had moderate SN for both modalities (radiograph SN = 0.50 and MRI SN = 0.54) and high SP (radiograph SP = 0.89 and MRI
SP = 0.94), resulting in high diagnostic utility for radiography (LR = 4.4) and even higher for MRI (LR = 9.4). There was no case for which the global impression was false-positive for both modalities.

The other primary results are the derived diagnostic values, assuming each imaging feature to individually represent a diagnostic test for the presence of SpA (Table 1). On radiography, sclerosis was seen in 77% of patients with JSpA, but also in half of controls, resulting in poor SP (0.44) and little diagnostic utility (LR = 1.4). On MRI, sclerosis was also seen frequently in the JSpA cohort (69% of patients) but with marginally better SP (0.66) and slightly better diagnostic utility (LR = 2.0). Erosion was frequently observed in patients with JSpA on both modalities (69% on radiography and 58% on MRI), but it occurred significantly less often in the control cohort (20% on radiograph and 9% on MRI, p < 0.001 in both modalities). The superior SP of erosion compared to sclerosis resulted in better diagnostic utility for this feature on both modalities (radiography LR ± 3.5) and erosion was especially diagnostic of JSpA on MRI (LR = 6.7). Joint space narrowing was the least common feature observed on both modalities (23% on radiography, 8% on MRI) but was marginally indicative of SpA when present on radiography (LR = 2.0), and slightly more so on MRI (LR = 2.7). Because of the inability to reliably evaluate joint space widening in preliminary testing, we did not assess joint space widening in our final reading exercise.

Regarding MRI-only features, BME was seen more frequently (62%) in the subjects with JSpA than was fat infiltration (38%), but BME was also observed in a surprisingly high proportion of controls (20%), whereas fat infiltration occurred in only 9% of controls. This resulted in good diagnostic utility for both features, but with higher positive LR for fat infiltration (LR = 4.5) than for BME (LR = 3.1). For patients with JSpA, fatty infiltration was significantly more common in those with longer disease duration.
(8/10 patients with disease > 2 yrs vs 6/16 with disease ≤ 2 years, p = 0.042), while the proportion of BME did not vary significantly by disease duration (4/10 vs 10/16, p = 0.024).

Results relating to interreader agreement are shown in Table 2. Discordant observation of sclerosis was common, resulting in only a fair amount of agreement on radiography (κ = 0.26) and MRI (κ = 0.25). Comparatively, results for erosion were more concordant on both radiography (κ = 0.48) and MRI (κ = 0.56), achieving a moderate level of agreement. Joint space narrowing was less common and reader agreement was especially poor for MRI (0.13). Agreement was generally good on the MRI-only features — substantial for marrow edema (κ = 0.62) and moderate for fat infiltration (κ = 0.50). Overall, global diagnostic impres-
sions formed on MRI had substantial agreement (κ = 0.80), while those formed on radiography had only fair agreement (κ = 0.30).

Reader 2 and the arbitrator were significantly more confident in their assignment of a positive diagnosis on MRI than radiography (3.8 vs 1.4 and 4.0 vs 2.8, p < 0.05). A similar trend seen for reader 1 did not reach statistical significance (2.9 vs 2.2, p > 0.05). The same pattern was observed for confidence in making a negative diagnosis.

The subanalyses we could perform were limited by small sample size. We compared the distribution of erosions by location on radiograph and MRI (Table 3). This demonstrated poor agreement between detection of erosions at the 2 modalities, except for substantial agreement at the right
iliac articular margin. Sample size was too small to meaningfully statistically compare findings between subgroups of patients with JSpA. Also, owing to the retrospective nature of this study, the radiographic projection could not be standardized. In the JSpA group, radiographs were either dedicated anteroposterior views of the SI joints (n = 14), or anteroposterior views of the pelvis, hips, or lumbar spine (n = 12). Although subjectively the readers agreed that dedicated films of the SI joints allowed improved diagnostic confidence, we found no significant numerical differences between findings stratified by film type.

DISCUSSION

Our study confirms the superiority of MRI over radiography for the diagnosis of JSpA affecting the SI joints. However, some of the results are surprising, with the superiority of MRI reflected mainly in improved reliability and SP compared to radiography, but with less overall SN for disease than some other studies have implied.

Prior studies have shown superiority of MRI over radiography for the diagnosis of JSpA. However, in our study, radiography performed better than has been recently suggested, and both imaging modalities detected only about...
half of JSpA cases, indicating that JSpA cannot be ruled out by negative imaging of the SI joints, including a negative MRI. This is consistent with the natural history of JSpA, where unlike in adult SpA, axial involvement of spine and SI joints often appears later in disease, with peripheral arthropathy and enthesopathy presenting earlier3,4,5. This observation is also consistent with a need for newer classification systems to develop criteria that accommodate the early peripheral patterns of disease16.

Sclerosis was observed frequently and was more frequently reported on radiography than MRI in both the disease and control groups. The diagnostic utility of the feature was poor owing to limited SP on modalities. Our findings are consistent with adult SpA literature, where “physiological” sclerosis is frequently observed17. Some have suggested that the SP of sclerosis as a finding in JSpA could be improved only by considering sclerosis abnormal if it extended 5 mm from the joint space8, but this threshold would have been highly insensitive in our pediatric cohort.

Compared to sclerosis, erosion was nearly as prevalent but was much more specific and was the single most useful diagnostic feature on both radiography (LR = 3.5) and MRI (LR = 6.7). Interreader reliability of erosion assessment was marginally better on MRI. The challenges involved in the interpretation of erosion are quite different between the 2 modalities. On radiographs, distinction from overlapping structures may create difficulty and the majority of the articular surfaces are not in profile. On MRI, the appearance of erosion is highly variable, with visualization often affected by MRI scan technique and by changes in signal of the surrounding bone marrow as a result of BME, fat infiltration, or sclerosis18. Poor visualization of the articular cortex of the SI joint on MRI is frequently observed in normal subjects, contributing to the challenge of interpretation. Reliability of detection of erosion appears to be slightly better in adults19, which is not surprising because assessment of the contours of the SI joint is more difficult in the immature skeleton. Defining all possible variations of the MRI appearance of erosion is challenging; however, with training, SI joint erosion in juveniles can be accurately observed on MRI. Most patients with erosions had multiple erosions; even if a particular erosion could not be detected on one imaging modality, the correct diagnosis could be made by detecting other erosions in the same patient. This explains how the overall diagnostic utility of erosion could be high even in the face of poor agreement between radiographic and MR identification of particular erosions at a given articular location. The diagnostic utility of erosion on MRI was more clinically useful despite controversies surrounding its definition (Figure 3).

Unlike in adults, where the natural history of SpA is progressive SI joint space loss and ultimately ankylosis, narrowing was uncommon in our pediatric JSpA cohort (23% on radiography, 8% on MRI), and no cases of ankylosis were observed. Comparatively, narrowing was slightly more diagnostically useful on MRI than on radiography, though far more reliably scored on radiography ($\kappa = 0.42$) than MRI ($\kappa = 0.13$). All readers noted that measures of joint space width (narrowing, widening, or ankylosis) would be of benefit in pediatric apophyseal joint assessment.

Figure 3. Erosion in a 14-year-old male with low back pain is easier to assess on magnetic resonance imaging (MRI) than on radiography, especially on the sacral side. A. Radiograph of the sacroiliac (SI) joint reveals bilateral iliac sclerosis with possible joint space widening, but no definite erosion is seen. B. One month later, MRI reveals clear evidence of erosion on this coronal T1-weighted image. Irregularity of the osteochondral interface is seen in several areas, with accompanying signal changes typical for active erosion best seen on both sides of the lower left SI joint. Changes on the right include sclerosis of the ilium with subtle erosion of the ilium inferiorly causing apparent joint space widening and marrow signal change in the right sacrum, without erosion on this image. The patient was subsequently diagnosed with reactive arthritis.
Bone marrow edema (BME) had moderate utility for diagnosis of JSpA and substantial interreader agreement. Other studies in both adult and juvenile SpA have shown higher sensitivity for inflammatory MRI changes, ranging from 0.81–1.00. However, these studies generally examined patients in followup settings in which a diagnosis of SpA had already been established prior to imaging. Studies more similar to ours, where imaging was performed at a time when patient diagnosis was still unknown, showed moderate sensitivity (SN = 0.54–0.67) more comparable to our results. While an important strength of BME is that it can be relatively reliably assessed, readers should be cautious about interpreting BME alone for the imaging diagnosis of JSpA because in this setting, 20% of control subjects revealed false-positive findings on the STIR sequence, although most of these findings were subtle. This confirms that the presence of BME alone may be insufficient for diagnosis, in contrast to recent recommendations in adults. Subjectively, we observed a variety of patterns of periarticular marrow signal at pediatric SI joints, not all pathologic. Further study could formally assess these patterns of BME in greater detail.

Subarticular fat infiltration on MRI was less sensitive for JSpA but highly specific when present. The overall diagnostic utility of this feature is thus good (LR = 4.5) and is in fact better than BME alone. The SN of fat infiltration is comparable to an adult study (SN = 0.38, SP = 0.81) comparing controls to patients with inflammatory back pain and preradiographic SpA, but the SP is better and this likely reflects a combination of factors including the presence of less overall physiological fat in the bone marrow of the pelvis in children. The significantly higher prevalence of fat infiltration in patients with disease duration > 2 years is in keeping with its nature as a finding of chronic disease.

Radiologist global impression for the diagnosis of JSpA was almost as accurate as erosion alone for both modalities, with erosion being a slightly more sensitive variable. However, global impression had higher diagnostic utility than erosion on radiography (LR = 4.4) and utility was especially high on MRI (LR = 9.4). Interreader reliability was highest of all for global impression of MRI (κ = 0.80) whereas when assessing a radiograph, erosion (κ = 0.48) had better reliability than global impression (κ = 0.30). Readers were also significantly more confident in their diagnoses on MRI. Global impression, integrating, and “averaging” multiple imaging features allowed more confident and consistent interpretation by our readers than any single feature.

Other studies of pediatric SI joints have reported imaging features such as erosion, sclerosis, and marrow edema without defining these features. One contribution of our study is to provide pediatric-specific definitions of these imaging features, developed by an iterative consensus process and evaluated in these data. These definitions can be applied to future studies to minimize interobserver variability and maximize diagnostic efficacy.

Our institution does not routinely use gadolinium contrast-enhanced MRI of SI joints. In some published studies of MRI of pediatric SI joints, gadolinium has been used, and in some it has not. Our current study shows the superiority of MRI over radiography, even without gadolinium-enhanced sequences. Given the methodological differences between these studies, it remains unclear whether or to what extent gadolinium enhancement would improve this performance. Given that the positive LR for global impression on MRI was 9.4, we do not expect significantly improved utility in our population from the addition of contrast-enhanced sequences.

In the 2 most recent pediatric studies, MRI showed findings in patients with JSpA who had essentially normal radiographs. Both had younger patients than our study (mean age 9.3 yrs and 12 yrs), and neither study performed systematic evaluation of radiographic findings or had a control group. All 11 patients in 1 study had MRI findings of sacroiliitis, which raises the question of circular reasoning and whether the positive MRI result was an indirect inclusion criterion. Sacroiliitis was seen in 29% of patients on MRI in the other study and it is clear that radiographs taken early in disease, before structural damage is established, are likely to be negative. In our study, patients had on average a 2-year history of symptoms when carefully questioned, but these MRI were performed for diagnosis of the etiology of low back pain/SI joint pain, rather than followup of known established disease. The results should therefore be relevant to patients initially presenting to a rheumatologist with symptoms suggesting possible SpA.

Our study has limitations. It is retrospective, and so radiographic and MRI protocols varied somewhat during the study period. The sample size was limited by the logistical difficulties of imaging young children, and there was a minor delay (4–5 mos) between radiography and MRI in most patients. However, structural damage in JSpA usually accumulates relatively slowly, reducing the effect of such delays on our findings. The results of our study should be applied only to pediatric patients who met our entry criteria, i.e., who have, in the opinion of an expert clinician, sufficient symptoms of low back/SI joint pain to be sent for MRI and radiography of the SI joints to determine etiology. Patients in other clinical settings, such as a new diagnosis of juvenile inflammatory arthritis with only involvement of peripheral joints/entheses, were not directly studied here.
and may have a different burden of disease. On demographic review of all patients who underwent MRI of the SI joints but did not have any radiographs, there was no significant difference in age or sex between that MRI-only group and the MRI plus radiography group (our study population).

MRI can reliably assess most features of sacroiliitis in children, and global impression of an MRI scan of the SI joints is especially reliable with high utility for the diagnosis of juvenile-onset SpA. While radiography is inexpensive, readily available, and has good diagnostic utility, MRI is the superior test in children with low back/SI joint pain that may be inflammatory, and radiography could be reserved for situations where access to MRI is limited.

REFERENCES