Assessment of sacroiliitis by radiographs and MRI: where are we now?

Gaalen, F.A. van; Bakker, P.A.C.; Hooge, M. de; Schoones, J.W.; Heijde, D. van der

Assessment of sacroiliitis by radiographs and MRI: where are we now?

Floris A. van Gaalen[a], *Pauline A.C. Bakker*[a], *Manouk de Hooge*[a], *Jan W. Schoones*[b], *and Désirée van der Heijde*[a]

Purpose of review
Both MRI and plain radiography are used to assess sacroiliitis. A weakness of radiography – apart from its inability to detect early disease – is reader variability. On the contrary, experience with MRI is relatively limited by comparison.

Recent findings
This review summarizes recent advances in sacroiliac joint imaging using radiography and MRI in spondyloarthritis.

Summary
Observer variation in reading radiographs of sacroiliac joints remains an unresolved issue. In recent years, more studies on MRI in the diagnosis of axial spondyloarthritis have become available. Incorporating structural lesions in the sacroiliac joint and spine and inflammatory lesions in the spine in the definition of a positive MRI are hot topics in research.

Keywords
ankylosing spondylitis, axial spondyloarthritis, diagnosis, MRI, radiography

INTRODUCTION

Sacroiliitis is the hallmark of axial spondyloarthritis (SpA) [1]. Axial SpA is called nonradiographic axial SpA when there are no (definite) abnormalities detected on plain film radiographs of the sacroiliac joint and is called radiographic axial SpA or ankylosing spondylitis (AS) when definite signs of sacroiliitis are seen on radiographs of the sacroiliac joints [2].

Radiography is the method most commonly used to assess involvement of the sacroiliac joint, but it is often inadequate to detect early disease, as patients may have symptoms for several years before abnormalities can be seen on radiography [3]. Moreover, reading radiographs of the sacroiliac joints are considered difficult. Interobserver and intraobserver variations are substantial, which implies that sacroiliitis is often missed or incorrectly diagnosed [4].

van Tubergen *et al.* [4] set out to investigate if training could improve performance of radiologists and rheumatologists in reading radiographs of the sacroiliac joints. One hundred rheumatologists and 23 radiologists took part in the study wherein sensitivity was assessed using sacroiliac joint radiographs of Human Leukocyte Antigen (HLA)-B27-positive AS patients, and specificity was assessed using radiographs of healthy HLA-B27-negative relatives. Participants scored radiographs at baseline, 3 months later after following a self-education program and again 3 months later after attending a workshop.

At baseline, median sensitivity of rheumatologists for detecting sacroiliitis on radiographs was 81% (range 31–100%) with a specificity of 75% (range 38–100%). After self-education, median sensitivity dropped to 75% (range 25–100%), whereas specificity increased to 78% (range 44–100%). After the workshop, the sensitivity then returned to 81% (range 25–100%), and specificity was stable at 79% (range 39–96%). The results of the radiologists showed similar fluctuations. At baseline, median sensitivity of radiologists for detecting sacroiliitis on radiographs was 88% (range 25–100%) with a specificity of 71% (range 46–100%). After self-education, median sensitivity dropped to 78% (range 44–100%), whereas specificity was 73% (range 38–96%). After the workshop, the sensitivity

[a]Department of Rheumatology, Leiden University Medical Center and [b]Walaeus Library, Leiden University Medical Center, Leiden, The Netherlands

Correspondence to Floris A. van Gaalen, MD, PhD, Department of Rheumatology, Leiden University Medical Center, C1-R, PO Box 9600, Leiden 2300 RC, The Netherlands. E-mail: f.a.van_gaalen@lumc.nl

Curr Opin Rheumatol 2014, 26:384–388
DOI:10.1097/BOR.0000000000000066

KEY POINTS

- Reader variability in reading radiographs of the sacroiliitis joints remains unresolved.
- As a result, there is concern about the reliability of radiographs of the sacroiliac joints in diagnosing axial SpA.
- Although more data are becoming available, the number of high-quality studies on the diagnostic utility of MRI of the sacroiliac joints is relatively limited.
- Incorporation of structural changes into the definition of a positive sacroiliac MRI is an important line of current research.

and specificity then increased to 84% (range 50–100%) and 85% (range 50–96%), respectively.

Intraobserver variation was tested using a set of 10 radiographs with various degrees of sacroiliitis. Agreement was high for radiographs without signs of sacroiliitis or complete ankylosis of sacroiliac joints (range of means 94–100%), but much lower for radiographs with more subtle changes (grade 1 or 2 sacroiliitis) with mean agreements ranging from 52 to 87%.

This study is important for several reasons. First, it shows that rheumatologists and radiologists have only modest sensitivity and specificity for diagnosing sacroiliitis using radiographs alone. In addition, there is considerable intraobserver variability particularly when changes because of sacroiliitis are subtle. Most importantly, however, this study shows that training of readers of radiographs does not improve their sensitivity or specificity. This means that, in its current form, there is concern about the reliability of conventional radiographs of sacroiliac joints when used for diagnosis.

MRI has proven capable of detecting inflammatory lesion in the sacroiliac joints in SpA before changes are seen on radiographs. An indication of the increased use of MRI in SpA is the inclusion of MRI of the sacroiliac joints in the Assessment of SpondyloArthritis international Society (ASAS) axial SpA classification criteria [5]. MRI of the sacroiliac joints is able to detect several features associated with sacroiliitis, such as ankylosis, bone marrow edema (BME)/osteitis, capsulitis, enthesitis, erosions, fat deposition and synovitis. However, a 2009 report by radiologists and rheumatologists from the ASAS/Outcome Measures in Rheumatology (OMERACT) MRI working group considered clear presence of BME essential for defining active sacroiliitis. In what has become known as the ASAS definition for a positive MRI, a BME lesion highly suggestive of sacroiliitis needs to be present in subchondralor periarticular bone. If there is only one signal (BME lesion) on an MRI slice, this BME lesion should be present on at least two consecutive slices, although when there are two or more BME lesions on a single slice, one slice is sufficient [6].

A systematic review published in 2012 reviewed the literature on MRI in SpA published until November 2011. The aim of the review was to determine the level of evidence for the utility of MRI in relation to the clinical diagnosis of SpA [7▪]. Studies included in the review had to be case–control or cohort studies and had to include the arbitrary number of more than 20 patients and 20 controls. After literature search, 76 full text articles were reviewed with only nine studies included in the review. Of these nine studies, only two met the authors' criteria for a high-quality report. Of these two reports, one reported on MRI abnormalities in the spine and one on MRI abnormalities in the sacroiliac joints. The authors of the review concluded that because of the small number of high-quality studies, current evidence for MRI in the diagnosis of axial SpA is limited.

With data showing that education does not improve reading radiographs of the sacroiliac joints and a review concluding that there are not enough studies demonstrating the diagnostic utility of MRI for sacroiliitis, we searched the literature for recent studies to see what progress has been made in recent years on the two most commonly used imaging modalities for detecting sacroiliitis.

LITERATURE SEARCH

PubMed, *Embase*, *Web of Science* and the *Cochrane Library* were searched in November of 2013 for articles on interobserver and intraobserver variability of radiographs of the sacroiliac joints in SpA and on the diagnostic value of MRI of the sacroiliac joints in SpA using two separate search strategies. The search strategy used for *PubMed* can be found in the supplementary Table 1, http://links.lww.com/COR/A15. All articles were reviewed by title and abstract by two out of three assessors (F.G., P.B. and M.H.), and an article was selected if both assessors agreed that the study contained data relevant to the search. The literature search for articles on radiographs was limited to publications after the study by Van Tubergen *et al.* [4] was published in 2003, and the search for MRI was limited to publication from 2010 onward, as the review by Arnbak *et al.* [7▪] had covered all articles published before 2011.

RECENT LITERATURE ON OBSERVER VARIATION IN READING RADIOGRAPHS OF THE SACROILIAC JOINTS

With the search strategy for radiographs, 800 articles were found in the databases. After review, five full

text articles and one meeting abstract were identified as containing relevant data.

In their 2004 study, Spoorenberg *et al.* [8] compared reliability and change over time of several radiological scoring methods in AS, including grading of sacroiliac joints using the 0–4 New York method and the almost identical Stoke Ankylosing Spondylitis Spine Score. Radiographs of 217 AS patients at baseline, 12 and 24 months were scored by two observers. Kappa values for intraobserver variability ranged from 0.36 to 0.76 and interobserver variability ranged from 0.66 to 0.70. Kappa coefficients are a statistical measure of interreader agreement that are thought to be more robust than simple agreement calculation as kappa takes agreement occurring by chance into account. Using the cutoff values proposed by Landis and Koch [9], in this study, the kappa for intraobserver variability indicates fair to substantial agreement and moderate-to-substantial agreement for interobserver variability. In the same year, another study [10] on radiological scoring methods for AS by the same group was published, but this did not provide separate results on variability of scoring sacroiliac joints radiographs.

An example of how variability in grading sacroiliac radiographs may affect clinical practice is provided by data from the multicenter German Spondyloarthritis Inception cohort (GESPIC) cohort [11]. Radiographs of the sacroiliac joints of 149 nonradiographic SpA and 182 AS patients were rescored by two central readers. After rescoring, 11.4% AS patients were reclassified into nonradiographic axial SpA, and 15.5% nonradiographic axial SpA were reclassified into AS. Agreement between the two readers was modest [intraclass correlation coefficient (ICC)] for the left sacroiliac joint, 0.36 [95% confidence interval (CI) 0.22–0.49] and for the right sacroiliac joint, ICC 0.36 (95% CI 0.22–0.49). In a more recent study [12] from the GESPIC cohort, kappa values for scoring the sacroiliac joints ranged from 0.51 to 0.59 between two readers indicating moderate agreement.

In a study assessing the performance of computed tomography (CT) of the sacroiliac joints in patients with suspected SpA, two radiologists independently read 100 paired radiographs and CT scans. Similarly to previous studies, interreader variability was moderate for sacroiliitis on radiographs (kappa 0.59) but was much better for CT scans of the sacroiliac joints (interobserver kappa 0.91) [13▪].

In summary, all studies published in the past 10 years confirm the substantial interobserver and intraobserver variability in grading radiographs of the sacroiliac joints for sacroiliitis, but we found no progress in solving or reducing the problem by education or technical innovation. A possible exception was a meeting abstract reporting slightly better performance of posterior-anterior projection as compared with anterior-posterior projection in radiographs of the sacroiliac joints, but the data have not yet been published [14].

RECENT LITERATURE ON THE DIAGNOSTIC UTILITY OF MRI FOR SACROILIITIS IN SPONDYLOARTHRITIS

With the search strategy for MRI, 1094 articles and meeting abstracts published since 2010 were found in the databases. After review, four full text articles were identified as containing relevant data.

Because of several months' overlap of the literature search, the first article found was the one high-quality study [7▪] described in the aforementioned systematic review. Weber *et al.* [15] published data from a cross-sectional, international multicenter study called MORPHO. Aim of the study was to assess the diagnostic utility of MRI in SpA and construct a definition for a positive MRI. After calibrating readers using a training set, five readers independently read MRI scans from 75 patients with AS, 27 patients with inflammatory back pain (IBP) suspected of having SpA, 26 patients with nonspecific back pain and 59 healthy controls. AS was diagnosed according to the modified New York criteria, IBP was defined by expert opinion, the Calin IBP criteria or the Berlin IBP criteria, and nonspecific back pain was defined on clinical grounds. For all MRIs, BME, erosions, fat infiltration and ankylosis were scored by the readers, and readers were asked if they thought the MRI scan confirmed the presence of SpA by global assessment. Using global assessment of the MRI, agreement for the diagnosis of SpA in IBP patients was 85% for all five readers and agreement for the absence of SpA was 92% in nonspecific back pain patients and 95% in healthy controls.

Comparing IBP patients with patients with nonspecific back pain and controls, the global assessment of the readers had a sensitivity of 51% and a specificity of 98%. The ASAS definition of a positive MRI using BME only had a sensitivity of 67% and a specificity of 88% and a new proposed definition based on BME and erosions (MORPHO definition) had a sensitivity of 81% and a specificity of 88%. Comparing AS patients with nonspecific back pain patients and controls, global assessment had a sensitivity of 90% and a specificity of 97%, and the ASAS definition had a sensitivity of 85% and a specificity of 88%.

In a study [16] from the same group, the same number of MRIs was scored by a different number of

readers, and this study showed that besides BME structural lesions, such as erosions, are commonly seen in AS and IBP although erosions were now more often scored in nonspecific back pain patients and controls than in their previous study. The authors subsequently published a study [17] using selected patients and controls from the previous two studies. MRIs of the sacroiliac joints of 30 AS patients and 30 controls were used to assess the reproducibility of scoring erosions using four readers. The kappa value for scoring erosions was 0.72, which was slightly higher than the kappa of 0.61 for scoring BME.

In their next article, Weber *et al.* [18▪] took a slightly different approach by using both the consensus classification of an MRI as SpA or no-SpA using global assessment and the clinical diagnosis as the gold standard for disease. In the consensus classification, an MRI was marked as consistent with SpA when at least three out of four readers thought the images showed signs of SpA. An MRI was marked as no-SpA when all readers thought that the images showed no signs of SpA with a high confidence.

MRIs of the sacroiliac joints from two inception cohorts were scored for presence of BME, erosions and fat infiltration by all four readers using a scoring system wherein the sacroiliac joint is represented by four quadrants (upper ilium, lower ilium, upper sacrum and lower sacrum). Cohort A consisted of 10 healthy controls and 79 patients of which 10 had AS, 20 nonradiographic axial SpA and 39 nonspecific back pain. Cohort B consisted of 88 patients with an acute uveitis and back pain who were referred to a rheumatologist. Diagnosis was made based on the clinical opinion of a rheumatologist. In this cohort, 31 patients had nonradiographic axial SpA, 24 AS and 33 patients with nonspecific back pain.

Using the consensus classification of sacroiliac joint MRI as the gold standard, to reach a preset specificity of 90%, BME had to be present in two sacroiliac joint quadrants. At this cutoff which the authors state is similar to the ASAS definition, sensitivity was 91% in cohort A and 83% in cohort B. To reach 90% specificity, only one erosion had to be present in both cohorts giving a perfect sensitivity of 100%. Using the clinical diagnosis as the standard, BME had to be present in three sacroiliac quadrants in cohort A and in four sacroiliac quadrants in cohort B to reach a specificity of 90%. At this cutoff, sensitivities were 73% for cohort A and 39% for cohort B. To reach 90% specificity, one erosion had to be present in cohort A and two erosions in cohort B, and this had a sensitivity of 77 and 54%, respectively. The combined features of BME and/or erosion had a sensitivity of 82% for cohort A and 51% for cohort B with a specificity of 90%. Irrespective of the standard used, fat infiltration performed worse than BME and erosions. The authors conclude that these results support the use of both BME and erosions in defining a positive MRI sacroiliac in axial SpA.

DISCUSSION

The literature of the past 10 years confirmed reader variability in reading radiographs of the sacroiliac joints, but no progress was made in reducing the variability. As the example of reclassification of AS and nonradiographic axial SpA patients by central readers in the GESPIC cohort exemplifies, physicians should be careful in making or rejecting the diagnosis of axial SpA based on radiographs of the sacroiliac joints alone.

Given the curved shape of the sacroiliac joint, which complicates radiography, CT has been investigated as an alternative imaging modality in suspected SpA [13▪] and AS [19]. However, a disadvantage of CT scans is that the radiation dose is higher than for radiographs. Given that axial SpA usually start in young adults: this is particularly an issue for young women in whom the ovaries are within the primary CT beam. Low radiation CT scanning protocols of the sacroiliac joints have been developed but are not in widespread use [20].

The literature on MRI of the sacroiliac joints in SpA of the last 3 years consisted of publications from a consortium from Switzerland, Denmark and Canada. One thing that is clearly encouraging about their data is that they showed that their expert readers had a good agreement on diagnosing SpA or no-SpA using MRI. In addition, another group reported substantial or almost perfect interreader variability in scoring MRI sacroiliac changes [21]. So, the available data indicate that MRI of the sacroiliac joints has an acceptable interreader variability.

Weber *et al.* advocate the incorporation of structural changes into the definition of a positive sacroiliac MRI for SpA, as this could improve sensitivity and specificity. However, more studies are needed, and a generally accepted definition of what constitutes structural changes consistent with SpA has yet to be decided.

Another possibility to improve the diagnostic utility of MRI scanning is to look for inflammatory or structural lesions in the spine. A consensus-based definition of a positive spinal MRI for inflammatory lesions (spondylitis) and structural changes (fat deposition) has been published [22], and spinal inflammation detected in the absence of inflammation in the sacroiliac joints has been observed in SpA patients [23], but more studies are needed.

The number of high-quality studies on the diagnostic utility of MRI of the sacroiliac joints remains relatively limited. However, following the publication of the ASAS axial SpA classification criteria, interest in the early stage of axial SpA has increased greatly in recent years. Therefore, it is to be expected that the number of studies on MRI imaging in axial SpA from cross-sectional studies, clinical trials and inception cohorts will increase in the coming years.

CONCLUSION

In axial SpA, no progress has been made in radiography of the sacroiliac joints in recent years, but there is steady progress in research on diagnostic utility and reliability of MRI of the sacroiliac joints.

Acknowledgements

None.

Conflicts of interest

There are no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- ■ of special interest
- ■■ of outstanding interest

1. Rudwaleit M, van der Heijde D, Khan MA, *et al.* How to diagnose axial spondyloarthritis early. Ann Rheum Dis 2004; 63:535–543.
2. Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: do we need new criteria? Arthritis Rheum 2005; 52:1000–1008.
3. Feldtkeller E, Bruckel J, Khan MA. Scientific contributions of ankylosing spondylitis patient advocacy groups. Curr Opin Rheumatol 2000; 12:239–247.
4. van Tubergen A, Heuft-Dorenbosch L, Schulpen G, *et al.* Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? Ann Rheum Dis 2003; 62:519–525.
5. Rudwaleit M, van der Heijde D, Landewe R, *et al.* The development of Assessment of SpondyloArthritis international society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009; 68:777–783.
6. Rudwaleit M, Jurik AG, Hermann KGA, *et al.* Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009; 68:1520–1527.
7. ■ Arnbak B, Leboeuf-Yde C, Jensen TS. A systematic critical review on MRI in spondyloarthritis. Arthritis Res Ther 2012; 14:R55.

A review of the literature on MRI for axial SpA.

8. Spoorenberg A, de VK, van der Linden S, *et al.* Radiological scoring methods in ankylosing spondylitis. Reliability and change over 1 and 2 years. J Rheumatol 2004; 31:125–132.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33:159–174.
10. Wanders AJ, Landewe RB, Spoorenberg A, *et al.* What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 2004; 50:2622–2632.
11. Rudwaleit M, Haibel H, Baraliakos X, *et al.* The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum 2009; 60:717–727.
12. Poddubnyy D, Rudwaleit M, Haibel H, *et al.* Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. Ann Rheum Dis 2011; 70:1369–1374.
13. ■ Devauchelle-Pensec V, D'Agostino MA, Marion J, *et al.* Computed tomography scanning facilitates the diagnosis of sacroiliitis in patients with suspected spondylarthritis: results of a prospective multicenter French cohort study. Arthritis Rheum 2012; 64:1412–1419.

Provides a clear example of observer variation in reading radiographs.

14. Muela Y, Casasola-Vargas JC, Pelaez-Ballestas I, Burgos-Vargas R. The performance of radiography (X-rays) of the sacroiliac joints (SIJ) in comparison with magnetic resonance imaging (MRI) studies in individuals with inflammatory back pain (IBP) identified in a community study. J Rheum 2011; Conference:1192.
15. Weber U, Lambert RGW, Ostergaard M, *et al.* The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. Arthritis Rheum 2010; 62:3048–3058.
16. Weber U, Lambert RGW, Pedersen SJ, *et al.* Assessment of structural lesions in sacroiliac joints enhances diagnostic utility of magnetic resonance imaging in early spondylarthritis. Arthritis Care Res 2010; 62:1763–1771.
17. Weber U, Pedersen SJ, Ostergaard M, *et al.* Can erosions on MRI of the sacroiliac joints be reliably detected in patients with ankylosing spondylitis?: a cross-sectional study. Arthritis Res Ther 2012; 14:R124.
18. ■ Weber U, Zubler V, Pedersen SJ, *et al.* Development and validation of an MRI reference criterion for defining a positive SIJ MRI in spondyloarthritis. Arthritis Care Res (Hoboken) 2012.

Two new patient cohorts with data on clinical performance of MRI in diagnosis of axial SpA.

19. Geijer M, Gothlin GG, Gothlin JH. The clinical utility of computed tomography compared to conventional radiography in diagnosing sacroiliitis: a retrospective study on 910 patients and literature review. J Rheumatol 2007; 34:1561–1565.
20. Jurik AG, Hansen J, Puhakka KB. Effective radiation dose from semicoronal CT of the sacroiliac joints in comparison with axial CT and conventional radiography. Eur Radiol 2002; 12:2820–2825.
21. Arnbak B, Jensen TS, Manniche C, *et al.* Spondyloarthritis-related and degenerative MRI changes in the axial skeleton: an inter- and intra-observer agreement study. BMC Musculoskelet Disord 2013; 14:274.
22. Hermann KG, Baraliakos X, van der Heijde DM, *et al.* Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. Ann Rheum Dis 2012; 71:1278–1288.
23. van der Heijde D, Sieper J, Maksymowych WP, *et al.* Spinal inflammation in the absence of SI joint inflammation on MRI in patients with active non-radiographic axial spondyloarthritis. Rheumatology (United Kingdom) 2013; Conference:i53.