RHEUMATOLOGY

Original article

Performance of magnetic resonance imaging in the diagnosis of axial spondyloarthritis: a systematic literature review

Alexis Jones[1,*], Timothy J. P. Bray[2,*], Peter Mandl ᴵᴰ[3], Margaret A. Hall-Craggs[2], Helena Marzo-Ortega[4,5] and Pedro M. Machado ᴵᴰ[1,6,7]

Abstract

Objectives. To summarize the evidence on the performance of MRI for the diagnosis of axial SpA.

Methods. This was a systematic literature review of all studies from January 2013 to March 2017 including adult patients with clinically suspected axial SpA undergoing MRI. Studies from a previously published systematic literature review up to January 2013 were also included.

Results. Thirty-one studies were included. Six studies demonstrated good sensitivity and specificity for SI joint (SIJ) bone marrow oedema (BMO). Specificity was increased by the presence of other structural lesions alongside BMO, particularly erosions or fat infiltration. Four studies addressed the utility of SIJ fat infiltration, finding good sensitivity but poor specificity. SIJ erosions showed good specificity in five studies. Studies addressing high T1 signal in the SIJ, fluid signal in the SIJ, ankylosis, sclerosis, capsulitis, backfill and vacuum phenomenon reported limited diagnostic value. In the spine, four studies reported moderate sensitivity and specificity for corner inflammatory lesions, and four reported poor sensitivity and specificity for spinal fat infiltration. Five studies evaluated the added value of spinal MRI over SIJ MRI alone, with variable results depending on the cohort. Six studies addressed the effect of acquisition parameters on diagnostic accuracy: fat-saturated T2-weighted imaging and short tau inversion recovery (STIR) imaging showed comparable utility in identifying BMO. Three studies showed that gadolinium was of minimal added value in the detection of BMO.

Conclusions. These results confirmed the diagnostic utility of MRI in axial SpA. Performance varied according to the characteristics of the cohort and the number and combination of MRI lesions considered.

Key words: spondyloarthritis, ankylosing spondylitis, sacroiliitis, magnetic resonance imaging, inflammation, diagnosis, diagnostic performance

CLINICAL SCIENCE

Rheumatology key messages

- This systematic review summarizes the evidence on the performance of MRI for the diagnosis of axial SpA.
- Results provide information for recommendations aiming to standardize practice around the use of MRI scans in axial SpA.

Introduction

SpA encompasses a group of immune-mediated inflammatory diseases characterized by axial inflammation, peripheral arthritis, enthesitis, dactylitis and extra-articular features such as psoriasis, uveitis and inflammatory bowel disease [1]. The Assessment of Spondyloarthritis

International Society (ASAS) simplified the classification of SpA by dividing the group into axial [2] and peripheral SpA [3]. Peripheral SpA refers to disease with predominantly peripheral features, while axial SpA (axSpA) describes patients with disease predominately affecting the axial skeleton. The ASAS classification criteria are not diagnostic criteria but aim to differentiate groups of patients within a disease spectrum, mainly for research

[1]Department of Rheumatology, University College London Hospitals NHS Foundation Trust, [2]Centre for Medical Imaging, University College London, London, UK, [3]Department of Rheumatology, Medical University of Vienna, Vienna, Austria, [4]NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, [5]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, [6]Department of Rheumatology, Northwick Park Hospital, London North West University Healthcare NHS Trust, London and [7]Centre for Rheumatology and MRC Centre for Neuromuscular Diseases, University College London, London, UK

purposes. The diagnosis of axSpA should be made by the clinician based on the combination of clinical, laboratory and imaging features, with MRI being one of the imaging components.

MRI is the only imaging technique capable of detecting both active (inflammatory) and chronic (structural) lesions as well as their anatomical distribution, contributing to the early diagnosis of axSpA. MRI correlates with histological findings in axSpA [4], is a predictor of response to therapy and can be used to monitor disease activity over time [1].

Despite the clear utility of MRI in axSpA, there remains inconsistency around its use in clinical practice. A recent survey of 269 radiologists in acute UK National Health Service trusts/health boards showed substantial variability in the use of paramagnetic contrast, sequence choice and anatomical coverage [5]. This survey found that only 75% of radiologists were aware of the term axSpA, and only 31% and 25% were aware of the ASAS definitions of positive MRI of the SI joints (SIJ) and spine, respectively, [5]. Despite being widely accepted as a key diagnostic marker, bone marrow oedema (BMO) was not used as a potential diagnostic feature of axSpA by 18% of radiologists [5].

The heterogeneity of MRI protocols and image interpretation is likely to cause inconsistency in the way that axSpA is diagnosed and may lead to missed or delays in diagnosis and inadequate or unnecessary treatment for patients. As such, there is a need to standardize the use of MRI and a consensus on how MRI lesions should be interpreted in relation to axSpA.

The aim of this systematic literature review (SLR) is to summarize the available evidence on the diagnostic utility of MRI in axSpA, including the significance of specific lesions, the influence of anatomical coverage and effect of acquisition parameters. The results of this SLR will be used to inform future consensus exercises regarding the use of MRI in axSpA.

Methods

Research questions

Members of a British Society of Spondyloarthritis (BRITSpA) MRI task force (nine musculoskeletal radiologists and nine rheumatologists with an interest in axSpA) proposed clinically relevant research questions (RQs) related to key aspects of the use of MRI in axSpA. Three final research questions (RQ1–3) were formulated and agreed upon by consensus (Table 1).

These questions were framed according to the Population, Intervention, Comparator, Outcome format [6], as detailed in supplementary Tables S1–S3, available at *Rheumatology* online. For all three questions, the population of interest consisted of adult patients ($\geqslant$18 years) with suspected and/or established axSpA, and the reference standard consisted of a clinical diagnosis of axSpA (optimal scenario) or global imaging criteria considered suggestive of axSpA (suboptimal scenario). The outcomes of interest were the sensitivity, specificity and likelihood ratios for the diagnosis of axSpA; for RQ2 and RQ3

TABLE 1 RQs generated by the BRITSpA working group

RQ1	Which lesion, or combination of lesions, is most sensitive and specific for the diagnosis of axSpA?
RQ2	How does the choice of anatomical region influence diagnostic performance?
RQ3	How do MRI acquisition parameters influence diagnostic performance?

RQ: research questions; BRITSpA: British Society for SpondyloArthritis.

additional endpoints including the prevalence of spinal inflammation in groups with and without SIJ inflammation and additional metrics relating to sequence performance (see Supplementary Material, available at *Rheumatology* online).

Study selection and data extraction

The SLR was conducted by two reviewers (T.J.P.B. and A.J.) under the guidance of the methodologist (P.M.M.). The search strategy (see Supplementary Material, available at *Rheumatology* online) from a previous EULAR systematic review, addressing the role of imaging in SpA, was adopted [7]. MEDLINE (1946), Embase (1974) and Cochrane (1993) databases were searched without language restrictions. We included all studies performed between January 2013 and March 2017, in addition to relevant studies selected from the previous EULAR SLR, which included all studies from the inception of the databases up to January 2013 [7]. Each reviewer screened titles and abstracts of all citations independently, and potentially relevant articles were reviewed in full text (supplementary Fig. S1, available at *Rheumatology* online). Papers fulfilling the inclusion criteria underwent full data extraction (supplementary Tables S4–S6, available at *Rheumatology* online) and were assessed for risk of bias (RoB) (supplementary Tables S7–S9, available at *Rheumatology* online). Both reviewers independently retrieved data using a predefined data extraction sheet. The following data were extracted: main characteristics of study (authors, journal and year of publication), study design, number of included patients (subdivided into axSpA patients and controls), reference standard, features of interest, technical factors relating to the acquisition (magnetic field strength, slice thicknesses, use of gadolinium, acquisition planes, spine coverage and sequence parameters), and the relevant outcome data. For studies addressing the effect of acquisition parameters (RQ3), we also recorded technical performance metrics including the contrast-to-noise ratio (supplementary Figs S2–S4, available at *Rheumatology* online).

Quality assessment

Each study was assessed independently for RoB by the same two reviewers who conducted the SLR (T.J.P.B. and A.J.) using the Quality Assessment of Diagnostic

Accuracy Studies-2 tool [8] (supplementary Tables S7–S9, available at *Rheumatology* online). This tool involves RoB assessment in four domains (patient selection, index test, reference standard, flow and timing); the first three domains are also assessed for applicability concerns, resulting in seven separate assessments for each study. Each assessment produced a rating of 'low', 'high' or 'unclear' (assigned scores of 0, 1 and 2 respectively). Discrepancies between reviewers regarding study selection, data extraction and RoB assessment were solved by discussion; a third reviewer (P.M.M.) was available in case no consensus could be achieved.

Results

Of the 8304 articles screened, 31 studies were finally included (supplementary Fig. S1, available at *Rheumatology* online). Twenty articles related to the diagnostic accuracy of specific lesions on MRI in the diagnosis of axSpA (RQ1) [9–28], five articles related to the influence of anatomical coverage on diagnostic performance (RQ2) and six related to the influence of acquisition parameters (RQ3).

Diagnostic accuracy

Sensitivity, specificity and likelihood ratios for each of the studies investigating diagnostic accuracy are shown in Table 2; results for BMO and combinations of BMO with other features are also shown graphically in Fig. 1. The main study and patient characteristics of these studies are summarized in supplementary Table S4, available at *Rheumatology* online, and details on the RoB assessment are described in supplementary Table S7, available at *Rheumatology* online.

Sacroiliac joints

Six studies investigated the diagnostic utility of BMO in the SIJ [11, 15, 17, 23, 27, 29] (Table 3). In general, these studies showed that BMO was the most sensitive individual lesion for the diagnosis of axSpA, although sensitivity (0.35–0.91) and specificity (0.75–0.90) estimates varied depending on the patient cohort, definition used for the reference standard and number of MRI lesions used to categorize the patients [11, 15, 17, 23, 27, 29].

Defining a reference standard for axSpA is challenging. Expert clinical opinion has limitations and is frequently made with knowledge of imaging results, leading to circular interpretation. Imaging standards fail to reflect the full clinical picture of axSpA, and there is a well-known delay from disease onset to radiographic changes. Weber *et al.* [11, 27] used clinical examination and plain radiography to identify those patients with axSpA. In their earlier study, Weber *et al.* [15] used a 'global assessment of MRI' to confirm a positive diagnosis of axSpA. Jans *et al.* [23] used the ASAS classification criteria as their reference standard in patients undergoing MRI with inflammatory back pain. Wick *et al.* [17] used a retrospective diagnosis of axSpA from clinical notes—it is unclear whether MRI had been used to make this diagnosis. Marzo-Ortega *et al.*

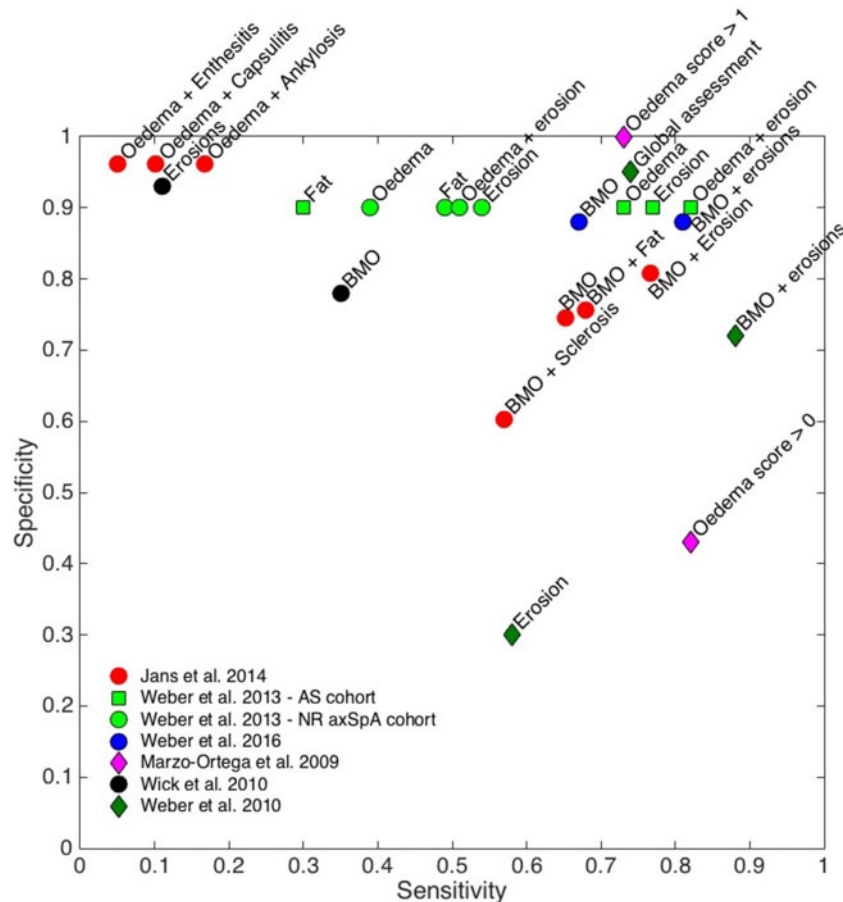[29] used Calin's criteria for the diagnosis of inflammatory back pain at baseline and 1 year.

There were subtle differences in the definition of BMO among authors. Jans *et al.* [23] defined a positive MRI SIJ for BMO if there was high fat-saturated T2-weighted (FS-T2W)/short tau inversion recovery (STIR) signal of the ilium or sacrum typically located periarticularly. If there was only one lesion, this had to be present on at least two consecutive slices. If there was more than one signal on a single slice, this was considered adequate. Weber *et al.* [11, 15] used a relatively similar definition using the Spondyloarthritis Research Consortium of Canada (SPARCC) assessment, where the SIJ is represented as a schematic with four quadrants. As with the ASAS definition, BMO had to be present in ⩾2 SIJs quadrants on the same slice or in the same SIJ quadrant on ⩾2 consecutive slices. In an earlier study, Weber *et al.* 2013 [27] used a cut-off of BMO in at least one quadrant. Marzo-Ortega *et al.* [29] used the Leeds scoring system: BMO was defined as low signal on T1 with enhancement after gadolinium administration and/or high or intermediate bone marrow signal with irregular contour on a T2 Spectral Presaturation with Inversion Recovery image. The presence of BMO was recorded and severity ranked on a semiquantitative scale based on the percentage area covered in each quadrant: 0, absent; grade 1, mild (<25%); grade 2, moderate (25–75%); grade 3, severe (75%). An overall score of inflammatory activity was calculated as the sum of scores of BMO. A positive MRI SIJ was defined as moderate/severe BMO (score ⩾2).

Both Jans *et al.* [23] and Weber *et al.* [11, 15, 27] found that the combinations of BMO and/or erosions could increase the sensitivity and specificity of MRI for the diagnosis of axSpA. Sensitivity and specificity were also increased by the combination of BMO and fat infiltration [21, 23]. Jans *et al.* [23] also reported an increase in specificity (but significant decrease in sensitivity) for the presence of BMO concomitantly with enthesitis, capsulitis or ankylosis.

Weber *et al.* investigated specific lesion-based criteria for defining a global positive sacroiliac joint MRI, and derived estimates of sensitivity and specificity for a number of different lesion cut-offs [11]. It was shown that lesion-based criteria including both BMO and erosions had superior sensitivity compared with criteria including BMO alone; for example, the presence of BMO in ⩾3 quadrants and erosions in ⩾3 quadrants produced sensitivity 0.83 and specificity 0.85 for the fulfilment of the global imaging criteria for axSpA [11]. However, estimates of sensitivity and specificity again varied substantially depending on the patient cohort.

Four studies addressed the utility of fat infiltration adjacent to the SIJ [10, 21, 23, 27]. The presence of fat infiltration was found to have low/moderate sensitivity (0.15–0.70) and moderate/high specificity (0.72–0.95) for the diagnosis of axSpA, although estimates varied depending on study design, the specific axSpA population under investigation and lesions' cut-offs [10, 21, 23, 27]. Weber *et al.* found that fat infiltration was more specific for

TABLE 2 Sensitivity and specificity of criteria using BMO and combinations in the sacroiliac joints

Feature	Study	Criterion	n	Reference standard	Sensitivity	Specificity	LR+	LR−
BMO	Jans et al. [23]	BMO	517	Clinical diagnosis axSpA	0.65	0.75	2.60	0.47
	Weber et al. [27]	BMO ⩾2 quadrants	177	Global MRI score	0.91/0.83[a]	0.90/0.90[a/b]	9.10/8.30[a]	0.10/0.19[a]
				Clinical diagnosis axSpA	0.73/0.39[a]	0.90/0.90[a/b]	7.30/3.90[a]	0.30/0.68[a]
	Weber et al. [11]	BMO ⩾2 quadrants (ASAS definition)	157	Clinical diagnosis axSpA	0.8/0.42[a]	0.76/0.73[a]	3.37/1.54[a]	0.26/0.80[a]
	Wick et al. [17]	BMO	179	Clinical diagnosis axSpA	0.35	0.78	1.59	0.83
	Weber et al. [15]	BMO ⩾2 quadrants (ASAS definition)	187	Clinical diagnosis axSpA	0.9	0.97	30.0	0.10
	Marzo-Ortega et al. [29]	BMO >0 (Leeds scoring)	76	Clinical diagnosis axSpA	0.82	0.42	1.41	0.43
BMO and erosions	Jans et al. [23]	BMO and erosion	517	Clinical diagnosis axSpA	0.77	0.81	4.05	0.28
	Weber et al. [27]	BMO and/or erosion ⩾1 quadrant	177	Global MRI score	0.98/0.96[a]	0.90/0.90[a/b]	9.80/9.60[a]	0.02/0.04[a]
				Clinical diagnosis axSpA	0.82/0.51[a]	0.90/0.90[a/b]	8.20/5.10[a]	0.20/0.54[a]
	Weber et al. [11]	BMO ⩾2 quadrants and ⩾1 erosion (MORPHO definition)	157	Clinical diagnosis axSpA	0.88	0.72	3.14	0.17
	Weber et al. [15]	BMO and erosion	187	Clinical diagnosis axSpA	0.9	0.97	30.0	0.10
BMO and fat infiltration	Jans et al. [23]	BMO and fat	517	Clinical diagnosis axSpA	0.68	0.76	2.83	0.42
	Weber et al. [21]	BMO and fat	157	Clinical diagnosis nr-axSpA	0.39	0.91	4.33	0.67
				Clinical diagnosis AS	0.58	0.91	6.44	0.46

[a]Values for two separate cohorts. [b]Pre-determined specificity. BMO: bone marrow oedema; LR: likelihood ratio; ASAS: Assessment of Spondyloarthritis International Society; nr-axSpA: non-radiographic axial spondyloarthritis.

Fig. 1 Diagnostic performance of bone marrow oedema and combinations in MRI SIJs

Sensitivity and specificity values are shown on a scatterplot for all relevant studies; performance for other features include in those studies (e.g. erosions alone) is also shown. SIJ: SI joint.

the diagnosis of AS than for non-radiographic axSpA (sensitivity/specificity 0.7/0.72 and 0.46/0.72, respectively) [21]. de Hooge *et al.* showed that using a cut off of ⩾3 fatty lesions correctly classified 63.6% of AS patients, whilst a combined threshold of ⩾5 fatty lesions and/or erosions performed similarly well [10].

Five studies investigated the diagnostic utility of erosions (Table 3) [10, 11, 17, 21, 27]. In general, erosions demonstrated good specificity for the diagnosis of axSpA, but only poor to moderate sensitivity (Table 3). Erosions were more sensitive in AS than in non-radiographic axSpA or axSpA as a whole [10, 21], and were more sensitive against a pre-specified MRI reference standard than against a clinical reference standard [27]. Using both erosions and fat infiltration as a diagnostic criterion increased specificity, but reduced sensitivity, compared with criteria consisting of fat infiltration alone [21].

Three studies addressing other SIJ lesions including high T1 signal in the SIJ, fluid signal in the SIJ, ankylosis, vacuum phenomenon, sclerosis, enthesitis, capsulitis and backfill reported low to moderate diagnostic performance for these features [9, 23, 24].

Spine

Five studies demonstrated moderate sensitivity and specificity of spinal inflammatory lesions in the diagnosis of axSpA (Table 4) [10, 12, 14, 22, 26]. In general, these studies demonstrated that lower thresholds for the number of inflammatory lesions resulted in reasonable sensitivity but poor specificity; increasing the threshold improved specificity but worsened sensitivity. Four of the five studies also investigated the diagnostic utility of spinal fatty lesions, and found poor sensitivity and high specificity, shown in Table 4 [10, 12, 22, 26]. de Hooge *et al.* [10], however, found ⩾5 spinal inflammatory lesions and ⩾5 spinal fatty lesions to be highly specific for axSpA, while still assuring an acceptable and useful level of discrimination between axSpA patients and non axSpA patients.

Effect of anatomical coverage

Five studies evaluated the added value of combined spinal and SIJ MRI over SIJ MRI alone [29, 30–33]. Study and patient characteristics for these studies are summarized in supplementary Table S5, available at

TABLE 3 Sensitivity and specificity of criteria using fat infiltration and erosions in the SIJs

Feature	Study	Criterion	n	Reference standard	Sensitivity	Specificity	LR+	LR-
Fat infiltration	de Hooge et al. [10]	Fat in ≥3 quadrants	287	Clinical diagnosis AS (mNY)	0.46	0.95[b]	9.20	0.57
				Clinical diagnosis nr-axSpA	0.15	0.95[b]	3.00	0.89
				Clinical diagnosis SpA (clinical arm)	0.15	0.95[b]	3.00	0.89
	Weber et al. [21]	Fat in ≥2 quadrants	157	Clinical diagnosis AS	0.70	0.73	2.59	0.41
				Clinical diagnosis nr-axSpA	0.44	0.73	1.63	0.77
	Jans et al. [23]	Presence of any fat	517	Clinical diagnosis SpA	0.55	0.84	3.44	0.54
	Weber et al. [27]	Lesion-based criteria for fat infiltration	177	Pre-specified positive MRI	0.34/0.74[a]	0.90/0.90[a]	3.40/7.40[a]	0.73/0.29[a]
				Clinical diagnosis axSpA	0.30/0.49[a]	0.90/0.90[a]	3.00/4.90	0.78/0.57
Erosions	de Hooge et al. [10]	≥3 erosions	287	Clinical diagnosis AS (mNY)	0.64	0.95[b]	12.80	0.38
				Clinical diagnosis nr-axSpA	0.47	0.95[b]	9.40	0.56
				Clinical diagnosis axSpA (clinical arm)	0.13	0.95[b]	2.60	0.92
	Weber et al. [11]	≥2 erosions	157	Clinical diagnosis axSpA	0.98/0.77	0.97/0.90	32.7/7.7	0.02/0.26
	Weber et al. [27]	Lesion-based erosion criteria	177	Pre-specified positive MRI	1/1[a]	0.90/0.90[a/b]	10	0
				Clinical diagnosis axSpA	0.77/0.54[a]	0.90/0.90[a/b]	7.70/5.40	0.26/0.51
Fat infiltration and erosions	Wick et al. [17]	Presence of any erosion	179	Clinical diagnosis axSpA	0.11	0.93	1.57	0.96
	Weber et al. [21]	Fat infiltration with erosion	157	Clinical diagnosis AS	0.68	0.98	34.00	0.33
				Clinical diagnosis nr-axSpA	0.34	0.98	17.00	0.67

[a]Values for two separate cohorts. [b]Pre-determined specificity. SIJ; SI joint; LR: likelihood ratio; mNY: modified New York criteria; SIJ: sacroiliac joint; nr-axSpA: non-radiographic axial spondyloarthritis.

TABLE 4 Sensitivity and specificity of criteria using inflammatory lesions and fatty lesions in the spine

Feature	Study	Criterion	n	Reference standard	Sensitivity	Specificity	LR+	LR−
Spinal inflammatory lesions	Weber et al. [26]	≥2 CILs	130	Clinical diagnosis axSpA	0.53/0.55[a]	0.64/0.74[a]	1.47/2.12[a]	0.73/0.61[a]
		≥3 CILs		Clinical diagnosis axSpA	0.43/0.25[a]	0.75/0.89[a]	1.72/2.27[a]	0.76/0.84[a]
	Weber et al. [14]	≥2 CILs	95	Clinical diagnosis axSpA	0.69	0.94	11.50	0.33
	de Hooge et al. [10]	Presence of spinal inflammatory lesions	287	Clinical diagnosis AS (mNY)	0.27	0.95[b]	5.40	0.77
				Clinical diagnosis nr-axSpA	0.14	0.95[b]	2.80	0.91
				Clinical diagnosis axSpA (clinical arm)	0.05	0.95[b]	1.00	1.00
	Hu et al. [22]	≥1 CIL	400	Clinical diagnosis AS (mNY)	0.52	0.55	1.16	0.87
		≥6 CILs			0.45	0.66	1.32	0.83
		≥11 CILs			0.04	0.78	0.18	1.23
	Bennett et al. [12]	≥1 inflammatory lesion[c]	185	Clinical diagnosis axSpA	0.67	0.56	1.52	0.59
		>3 inflammatory lesions			0.45	0.81	2.37	0.68
		>3 inflammatory lesions and age <50			0.33	0.97	11.00	0.69
Spinal fatty lesions	Weber et al. [26]	≥6 spinal fatty lesions	130	Clinical diagnosis axSpA	0.26/0.40[a]	0.82/0.81[a]	1.44/2.11[a]	0.90/0.74[a]
	de Hooge et al. [10]	Presence of spinal fatty lesions	287	Clinical diagnosis AS (mNY)	0.18	0.95[b]	3.60	0.86
				Clinical diagnosis nr-axSpA	0.22	0.95[b]	4.40	0.82
				Clinical diagnosis axSpA (clinical arm)	0.02	0.95[b]	0.40	1.03
	Hu et al. [22]	≥1 spinal fatty lesion	400	Clinical diagnosis AS (mNY)	0.13	0.94	2.17	0.93
		≥2 spinal fatty lesions			0.09	0.99	9.00	0.92
	Bennett et al. [12]	≥1 spinal fatty lesion[c]	185	Clinical diagnosis axSpA	0.33	0.93	4.71	0.72
		>3 spinal fatty lesions[c]			0.22	0.98	11.00	0.80

[a]Values for two separate cohorts. [b]Pre-determined specificity. [c]These lesions were referred to as Romanus lesions and fatty Romanus lesions in [11]. CIL: corner inflammatory lesions; LR: likelihood ratio; mNY: modified New York; nr-axSpA: non-radiographic axial spondyloarthritis.

Rheumatology online, and details on the RoB assessment are described in supplementary Table S8, available at *Rheumatology* online. Two studies found that combined spinal and SIJ MRI did not add significant value over SIJ MRI alone, either because spinal inflammation was rare in the absence of SIJ inflammation [30] or because combined MRI resulted in a high rate of false positives [31]. However, three studies observed spinal inflammation in up to half of patients without SIJ inflammation, arguing that combined MRI adds value over SIJ alone [29, 32, 33].

Effect of acquisition parameters on diagnostic performance

The key acquisition parameters including method of fat suppression, anatomical coverage and use of contrast for all included studies, are summarized in supplementary Figure S2, S3 and S4, available at *Rheumatology* online. Six of the included studies specifically investigated the effect of acquisition parameters [34–39]. Study and patient characteristics for these studies are summarized in supplementary Table S6, available at *Rheumatology* online, and details on the RoB assessment are described in supplementary Table S9, available at *Rheumatology* online.

Of the six studies, three investigated the effect of sequence choice on diagnostic accuracy of axSpA or on the characteristics of the images themselves [34–36]. Boy *et al.* found that sensitivity and specificity were highest for FS-T2W imaging, and progressively decreased for STIR, diffusion-weighted and dynamic-contrast enhanced images, respectively [34]. Dalto *et al.* showed good levels of agreement between FS-T2W imaging and STIR imaging, with a Lin's concordance correlation coefficient of 0.94 for reader 1 and 0.88 for reader 2 (range 0–1) [36]. Ozgen investigated the role of T2-weighted Dixon imaging in the identification of BMO, and found a superior contrast-to-noise ratio compared with FS-T2W imaging [35]. Three studies investigated the role of gadolinium in the SIJs, and overall found minimal or no added value [37–39].

Discussion

We systematically reviewed the literature regarding the use of MRI in the diagnosis of axSpA, informing a task force of radiologists and rheumatologists with the aim of standardizing the use of MRI in suspected axSpA.

Overall, studies investigating specific SIJ MRI lesions have shown that BMO is the most sensitive and specific individual lesion. Structural lesions including fat infiltration have moderate sensitivity and specificity, whilst erosions demonstrate good specificity but relatively poor sensitivity. An important consideration is that several of these studies use fixed specificity values; it is likely that specificity would be lower, but sensitivity higher, if these values were allowed to vary freely.

Other SIJ lesions including high T1 signal in the SIJ, fluid signal in the SIJ, ankylosis, vacuum phenomenon, sclerosis, enthesitis, capsulitis and backfill have a low to moderate diagnostic utility, and are, therefore, unlikely to be of diagnostic value in isolation. Owing to the heterogeneity of the data, with varying reference standards and patient

cohorts across studies, or repeated use of the same cohort (implying an overlap in at least part of the study populations) we have been unable to create an accurate meta-analysis of lesion-based criteria in the diagnosis of axSpA.

A number of studies have assessed combinations of lesions and their diagnostic performance. These studies showed that a combination of BMO and erosions, or BMO and fat infiltration, yielded higher sensitivity and specificity than BMO alone. Pre-defined numbers of lesions or cut-offs have also been analysed and suggest that BMO in ⩾3 quadrants and erosions in ⩾3 quadrants show high sensitivity and specificity, and presence of 3–5 fatty lesions also yields good sensitivity. However, further studies are required to validate these findings.

In the spine, studies investigating the value of spinal inflammatory lesions found moderate sensitivity and specificity, whilst spinal fatty lesions were found to have relatively poor sensitivity and specificity. Although the results suggest that spinal lesions alone are unlikely to have sufficient diagnostic performance for use in axSpA, these lesions might be useful in combination with features identified on SIJ MRI. This is an area that requires further research.

The results of studies investigating the effect of anatomical coverage on diagnosis were mixed: two studies suggested that spinal inflammation is rare in the absence of SIJ inflammation, while three found the opposite. Assuming patients seen in clinical practice have variable presentations, imaging the spine would facilitate the diagnosis and management of patients with axial pain. Unfortunately, even amongst studies that have imaged the spine, there has been substantial heterogeneity in anatomical coverage (Supplementary Material, available at *Rheumatology* online) and there is clearly scope for further work to determine the 'optimal' spinal protocol. Importantly, this research will need to consider the trade-off between scan time (and therefore also cost) and diagnostic yield, particularly as pressures on radiology departments continue to increase.

The number of studies assessing the impact of acquisition parameters on diagnostic accuracy was relatively small. The available evidence suggests that contrast adds little value, although no studies have rigorously addressed this question in the spine. Again, there is a need for further research to address this issue.

Of the studies specifically investigating sequence choice, several studies investigated methods of fat suppression other than STIR imaging. FS-T2W was shown to have superior sensitivity and specificity to STIR imaging [34], with assessments of disease severity at the MRI level agreeing closely between the two sequences [36]. Similarly, Ozgen [35] demonstrated superior contrast-to-noise ratios for T2W Dixon imaging compared with STIR, but did not assess diagnostic sensitivity. Overall, these methods are promising alternatives to STIR and may offer improvements in image quality in the future.

There are several limitations of the studies included in this SLR. First, a number of the studies were potentially

biased by the inclusion of information from MRI scans in their reference standard (supplementary Tables S7–S9, available at *Rheumatology* online). In some studies, a positive MRI scan was used as an inclusion criterion; other studies selected patients based on previous MRI scans. Even in those studies that did not explicitly use MRI-based reference standards, it is unclear whether MRI had been used in the patients' prior diagnostic work-up or referral.

A true assessment of the diagnostic utility of MRI would omit any MRI imaging from the reference standard. However, in the absence of a robust biomarker for the disease, finding an accurate and reliable reference standard poses a challenge. Some studies incorporated a purely clinical reference standard with a diagnosis of axSpA made by a panel of expert physicians. An alternative approach might be to use reference standards based on follow-up and assessment at multiple time-points, to ensure a high level of confidence in the diagnosis of axSpA.

The use of control groups by the included studies was suboptimal, resulting in 'unclear' or 'high' RoB for a number of studies when assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool. Healthy controls can artificially inflate the sensitivity and specificity statistics, since it is typically easier to distinguish axSpA from healthy patients than from patients with other axial problems, namely chronic non-specific low back pain.

On a similar note, there remains uncertainty about the frequency of MRI lesions in the general population. Marzo-Ortega *et al.* [29] reported a high prevalence of BMO affecting the SIJs of up to 6/22 (27%) in a control sample of healthy volunteers and patients with mechanical back pain. Similar findings have been found in post-partum women [40], runners [40], soldiers [41] athletes [42] and the general population [43]. In patients with chronic low back pain recruited from primary care without previous rheumatological assessment, 21% met the MRI classification criteria based on SIJ BMO alone, but 42% of these lesions were small and of questionable clinical relevance as they showed no association with clinical SpA features [44].

A further limitation of this SLR is that the numbers of studies included under each of the RQs was relatively small. The number of studies was particularly small for RQ2 and RQ3, however, all relevant papers available have been included. Further work is needed to answer these questions more definitively.

Future research into the use of MRI in axSpA should assess MRI scans longitudinally in a cohort of patients with suspected axSpA, correlating lesions with symptoms, response to treatment and rate of radiographic progression. This cohort should cover the entire spectrum of axial disease. Separate studies on healthy controls should aim to assess the background noise of SIJ and spinal lesions associated with mechanical causes in a normal population, providing guidelines on minimum requirements or 'cut-offs' for lesions to determine an abnormal scan.

Advanced MRI techniques, including quantitative MRI hold promise for more accurate assessment of inflammation. Quantitative MRI techniques typically use a succession of scans to 'probe' tissue characteristics, and infer attributes such as cellularity, vascularity or fat content. Each pixel (picture element) in a quantitative MRI image has a measurable numerical value that reflects the intrinsic properties of a tissue, rather than arbitrary signal intensity produced by standard MRI [45, 46]. The application of these techniques to axSpA could potentially improve the understanding and management of the disease both through improvements in precision and through a more detailed assessment of bone marrow pathophysiology.

To conclude, the results of this SLR have informed the recommendations of a consensus group aiming to standardize practice around the use of MRI scan in the UK [47] and can inform similar exercises in other countries or at the international level.

Acknowledgements

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet 2017;390:73–84.

2 Rudwaleit M, van der Heijde D, Landewé 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;68:777–83.

3 Rudwaleit M, van der Heijde D, Landewé R *et al*. The Assessment of SpondyloArthritis international Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25–31.

4   Puhakka KB, Melsen F, Jurik AG *et al*. MR imaging of the normal sacroiliac joint with correlation to histology. Skeletal Radiol 2004;33:15–28.

5   Bennett AN, Marzo-Ortega H, Kaur-Papadakis D, Rehman A. The use of magnetic resonance imaging in axial spondyloarthritis: time to bridge the gap between radiologists and rheumatologists. J Rheumatol 2017;44:780–5.

6   Sackett DL, Straus SE, Richardson WS, Rosenberg WMC, Haynes RB, eds. Evidence-based medicine: how to teach and practice EBM. 2nd ed. Philadelphia: Churchill Livingstone, 2000.

7   Mandl P, Navarro-Compan V, Terslev L *et al*. EULAR recommendations for the use of imaging in the diagnosis and management of spondyloarthritis in clinical practice. Ann Rheum Dis 2015;74:1327–39.

8   Whiting PF, Rutjes AWS, Westwood ME *et al*. Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–36.

9   Laloo F, Herregods N, Varkas G *et al*. MR signal in the sacroiliac joint space in spondyloarthritis: a new sign. Eur Radiol 2017;27:2024–30.

10 de Hooge M, van den Berg R, Navarro-Compán V *et al*. Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? Ann Rheum Dis 2016;75:1308–14.

11 Weber U, Østergaard M, Lambert RGW *et al*. Candidate lesion-based criteria for defining a positive sacroiliac joint MRI in two cohorts of patients with axial spondyloarthritis. Ann Rheum Dis 2015;74:1976–82.

12 Bennett AN, Rehman A, Hensor EMA *et al*. Evaluation of the diagnostic utility of spinal magnetic resonance imaging in axial spondylarthritis. Arthritis Rheum 2009;60:1331–41.

13 Bennett AN, Rehman A, Hensor EMA *et al*. The fatty Romanus lesion: a non-inflammatory spinal MRI lesion specific for axial spondyloarthropathy. Ann Rheum Dis 2010;69:891–4.

14 Weber U, Hodler J, Kubik RA *et al*. Sensitivity and specificity of spinal inflammatory lesions assessed by whole-body magnetic resonance imaging in patients with ankylosing spondylitis or recent-onset inflammatory back pain. Arthritis Rheum 2009;61:900–8.

15 Weber U, Lambert RGW, Østergaard 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;62:3048–58.

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 (Hoboken) 2010;62:1763–71.

17 Wick MC, Weiss RJ, Jaschke W, Klauser AS. Erosions are the most relevant magnetic resonance imaging features in quantification of sacroiliac joints in ankylosing spondylitis. J Rheumatol 2010;37:622–7.

18 Na RK, Choi JY, Sung HH *et al*. ''MR corner sign'': value for predicting presence of ankylosing spondylitis. Am J Roentgenol 2008;191:124–8.

19 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.

20 Jacquemin C, Rubio Vargas R, van den Berg R *et al*. What is the reliability of non-trained investigators in recognising structural MRI lesions of sacroiliac joints in patients with recent inflammatory back pain? Results of the DESIR cohort. RMD Open 2016;2:e000303.

21 Weber U, Pedersen SJ, Zubler V *et al*. Fat infiltration on magnetic resonance imaging of the sacroiliac joints has limited diagnostic utility in nonradiographic axial spondyloarthritis. J Rheumatol 2014;41:75–83.

22 Hu Z, Xu M, Lin Z *et al*. Limitations of lumbar spine MRI in the diagnosis of ankylosing spondylitis. Clin Exp Rheumatol 2014;32:22–7.

23 Jans L, Coeman L, Van Praet L *et al*. How sensitive and specific are MRI features of sacroiliitis for diagnosis of spondyloarthritis in patients with inflammatory back pain? JBR-BTR 2014;97:202–5.

24 Hu Z, Wang X, Qi J *et al*. Backfill is a specific sign of axial spondyloarthritis seen on MRI. Joint Bone Spine 2016;83:179–83.

25 Weber U, Maksymowych WP, Chan SM *et al*. Does evaluation of the ligamentous compartment enhance diagnostic utility of sacroiliac joint MRI in axial spondyloarthritis? Arthritis Res Ther 2015;17:246.

26 Weber U, Zhao Z, Rufibach K *et al*. Diagnostic utility of candidate definitions for demonstrating axial spondyloarthritis on magnetic resonance imaging of the spine. Arthritis Rheumatol 2015;67:924–33.

27 Weber U, Zubler V, Pedersen SJ *et al*. Development and validation of a magnetic resonance imaging reference criterion for defining a positive sacroiliac joint magnetic resonance imaging finding in spondyloarthritis. Arthritis Care Res (Hoboken) 2013;65:977–85.

28 Larbi A, Viala PO, Molinari N *et al*. Assessment of MRI abnormalities of the sacroiliac joints and their ability to predict axial spondyloarthritis: A retrospective pilot study on 110 patients. Skeletal Radiol 2013;43: 351–8.

29 Marzo-Ortega H, McGonagle D, O'Connor P *et al*. Baseline and 1-year magnetic resonance imaging of the sacroiliac joint and lumbar spine in very early inflammatory back pain. Relationship between symptoms, HLA-B27 and disease extent and persistence. Ann Rheum Dis 2009;68:1721–7.

30 Althoff CE, Sieper J, Song I-H *et al*. Active inflammation and structural change in early active axial spondyloarthritis as detected by whole-body MRI. Ann Rheum Dis 2013;72:967–73.

31 Weber U, Zubler V, Zhao Z *et al*. Does spinal MRI add incremental diagnostic value to MRI of the sacroiliac joints alone in patients with non-radiographic axial spondyloarthritis? Ann Rheum Dis 2015;74:985–92.

32 van der Heijde D, Sieper J, Maksymowych WP *et al*. Spinal inflammation in the absence of sacroiliac joint inflammation on magnetic resonance imaging in patients with active nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2014;66:667–73.

33 Larbi A, Fourneret B, Lukas C *et al.* Prevalence and topographic distribution of spinal inflammation on MR imaging in patients recently diagnosed with axial spondyloarthritis. Diagn Interv Imaging 2017;98:347–53.

34 Boy FN, Kayhan A, Karakas HM *et al.* The role of multiparametric MR imaging in the detection of early inflammatory sacroiliitis according to ASAS criteria. Eur J Radiol 2014;83:989–96.

35 Ozgen A. The value of the T2-weighted multipoint dixon sequence in MRI of sacroiliac joints for the diagnosis of active and chronic sacroiliitis. Am J Roentgenol 2017;208:603–8.

36 Dalto VF, Assad RL, Crema MD, Louzada-Junior P, Nogueira-Barbosa MH. MRI assessment of bone marrow oedema in the sacroiliac joints of patients with spondyloarthritis: is the SPAIR T2w technique comparable to STIR? Eur Radiol 2017;27:1–8.

37 Hermann KGA, Landewé RBM, Braun J, Van Der Heijde D. Magnetic resonance imaging of inflammatory lesions in the spine in ankylosing spondylitis clinical trials: is paramagnetic contrast medium necessary? J Rheumatol 2005;32:2056–60.

38 van Onna M, van Tubergen A, van der Heijde D, Jurik AG, Landewé R. Gadolinium contrast-enhanced MRI sequence does not have an incremental value in the assessment of sacroiliitis in patients with early inflammatory back pain by using MRI in combination with pelvic radiographs: a 2-year follow-up study. Clin Exp Rheumatol 2014;32:225–30.

39 de Hooge M, van den Berg R, Navarro-Compan V *et al.* Magnetic resonance imaging of the sacroiliac joints in the early detection of spondyloarthritis: no added value of gadolinium compared with short tau inversion recovery sequence. Rheumatology (Oxford) 2013;52:1220–4.

40 de Winter J, de Hooge M, van de Sande M *et al.* Magnetic resonance imaging of the sacroiliac joints indicating sacroiliitis according to the Assessment of SpondyloArthritis international Society definition in healthy individuals, runners, and women with postpartum back pain. Arthritis Rheumatol 2018;70:1042–8.

41 Varkas G, De Hooge M, Renson T *et al.* Effect of mechanical stress on magnetic resonance imaging of the sacroiliac joints: assessment of military recruits by magnetic resonance imaging study. Rheumatology (Oxford) 2018;57:508–13.

42 Weber U, Jurik AG, Zejden A *et al.* Frequency and anatomic distribution of magnetic resonance imaging features in the sacroiliac joints of young athletes: exploring ''background noise'' toward a data-driven definition of sacroiliitis in early spondyloarthritis. Arthritis Rheumatol 2017;70:736–45.

43 Baraliakos X, Feldmann D, Ott A *et al.* OP0243 prevalence of inflammatory and chronic changes suggestive of axial spondyloarthritis in magnetic resonance images of the axial skeleton in individuals < 45 years in the general population as part of a large community study (SHIP). Ann Rheum Dis 2018;77(Suppl 2):170–170.

44 Arnbak B, Grethe Jurik A, Hørslev-Petersen K *et al.* Associations between spondyloarthritis features and magnetic resonance imaging findings: a cross-sectional analysis of 1, 020 patients with persistent low back pain. Arthritis Rheumatol 2016;68:892–900.

45 Bray TJP, Vendhan K, Ambrose N *et al.* Diffusion-weighted imaging is a sensitive biomarker of response to biologic therapy in enthesitis-related arthritis. Rheumatology (Oxford) 2017;56:399–407.

46 Bray TJP, Bainbridge A, Punwani S, Ioannou Y, Hall-Craggs MA. Simultaneous quantification of bone edema/adiposity and structure in inflamed bone using chemical shift-encoded MRI in spondyloarthritis. Magn Reson Med 2018;79:1031–42.

47 Bray TJP, Jones A, Bennett AN *et al.* Recommendations for acquisition and interpretation of MRI of the spine and sacroiliac joints in the diagnosis of axial spondyloarthritis in the UK. Rheumatology 2019;58:1831–8.