











ORIGINAL RESEARCH

Learning imaging in axial
spondyloarthritis: more than just a
matter of experience

Felix Radny,¹ Katharina Ziegeler ,^{1,2} Iris Eshed ,^{3,4} Juliane Greese,¹ Dominik Deppe,¹ Carsten Stelbrink,¹ Robert Biesen ,⁵ Hiltrun Haibel,⁶ Valeria Rios Rodriguez,⁶ Judith Rademacher ,^{6,7} Mikhail Protopopov ,⁶ Fabian Proft ,⁶ Kay Geert A Hermann ,¹ Denis Poddubnyy ,⁶ Torsten Diekhoff ,^{1,7} Sevtap Tugce Ulas ,^{1,7}

To cite: Radny F, Ziegeler K, Eshed I, *et al.* Learning imaging in axial spondyloarthritis: more than just a matter of experience. *RMD Open* 2024;**10**:e003944. doi:10.1136/rmdopen-2023-003944

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/rmdopen-2023-003944>).

Received 25 November 2023
Accepted 19 February 2024

ABSTRACT

Objective Reliable interpretation of imaging findings is essential for the diagnosis of axial spondyloarthritis (axSpA) and requires a high level of experience. We investigated experience-dependent differences in diagnostic accuracies using X-ray (XR), MRI and CT.

Methods This post hoc analysis included 163 subjects with low back pain. Eighty-nine patients had axSpA, and 74 patients had other conditions (mechanical, degenerative or non-specific low back pain). Final diagnoses were established by an experienced rheumatologist before the reading sessions. Nine blinded readers (divided into three groups with different levels of experience) scored the XR, CT and MRI of the sacroiliac joints for the presence versus absence of axSpA. Parameters for diagnostic performance were calculated using contingency tables. Differences in diagnostic performance between the reader groups were assessed using the McNemar test. Inter-rater reliability was assessed using Fleiss kappa.

Results Diagnostic performance was highest for the most experienced reader group, except for XR. In the inexperienced and semi-experienced group, diagnostic performance was highest for CT&MRI (78.5% and 85.3%, respectively). In the experienced group, MRI showed the highest performance (85.9%). The greatest difference in diagnostic performance was found for MRI between the inexperienced and experienced group (76.1% vs 85.9%, $p=0.001$). Inter-rater agreement was best for CT in the experienced group with $\kappa=0.87$.

Conclusion Differences exist in the learnability of the imaging modalities for axSpA diagnosis. MRI requires more experience, while CT is more suitable for inexperienced radiologists. However, diagnosis relies on both clinical and imaging information.

INTRODUCTION

Imaging is of great importance in the diagnosis of axial spondyloarthritis (axSpA) as well as in therapy monitoring.^{1 2} Patients typically present with low back pain in young adulthood,³ but clinically it may be

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The diagnostic accuracy in assessing imaging findings depends on the radiologists' or rheumatologists' experience in axial spondyloarthritis (axSpA) imaging.

WHAT THIS STUDY ADDS

⇒ The level of experience and the modality has an impact on diagnostic accuracy.
⇒ Learnability differs across imaging modalities, with MRI being the most and X-ray the least teachable modality.
⇒ CT is the most accessible modality for less experienced readers.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Continuous training of young radiologists and rheumatologists in the field of axSpA imaging is crucial for accurate early diagnosis of axSpA and, thus, for improved patients' outcomes.
⇒ Especially in MRI, training can provide significant positive effects on diagnostic accuracy.

difficult to differentiate from other causes of back pain.^{4 5} The disease primarily affects the sacroiliac joints (SIJs) and the spine.⁶ If untreated, axSpA may lead to a loss of mobility and a significant worsening of the quality of life.⁷ For this reason, the early and reliable detection of axSpA with imaging techniques is crucial for diagnosis.⁸ As difficult as the diagnosis of axSpA is clinically,⁹ it is also challenging with imaging,¹⁰ as imaging parameters may be either inconclusive or prone to misinterpretation, especially in the early phase of disease.¹¹⁻¹³

In clinical routine, an X-ray (XR) examination of the SIJs is usually performed as the initial imaging test to detect structural



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Sevtap Tugce Ulas;
sevtap-tugce.ulas@charite.de

changes, such as erosions and ankyloses, associated with axSpA.¹ However, structural lesions are absent or undetectable with this imaging modality in non-radiographic axSpA,¹¹ and radiographs of the SIJs are notoriously difficult to interpret due to the complex anatomy of the joints. Reliable detection of inflammation, such as bone marrow oedema and other lesions, in early disease is only possible with MRI.¹⁴ Moreover, knowledge of key differential diagnoses, such as mechanical or degenerative SIJ disease, is also essential in evaluating these images since bone marrow oedema can also occur in mechanically induced conditions of the SIJs.¹⁵ In this respect, it is noteworthy that CT, the gold standard for detecting structural lesions, has recently gained attention as an imaging modality for patients with inconclusive MRI findings or contraindications to MRI or when MRI is not available.^{16–19}

The interpretation of imaging findings in axSpA requires a high level of radiological experience, which is one of the factors that contributes to the notoriously long delay between the onset of symptoms and diagnosis.²⁰ On the other hand, the uncritical use of classification criteria for diagnosis and misinterpretation of magnetic resonance images lead to a considerable number of false-positive diagnoses in current clinical practice.¹² Accessibility, costs, patient burden, diagnostic accuracy and reliability of imaging tests are critical factors for the decision, which imaging test to choose during a specific step of the diagnostic process.^{19 21}

This study aims to investigate the diagnostic accuracy in axSpA imaging of readers with different experience levels and to identify differences in their interpretation of XR, MRI and CT examination in order to elucidate which imaging modality or combination of modalities yield the highest diagnostic accuracy and improve reliability.

MATERIALS AND METHODS

Subjects

This study included two different prospective populations of patients with chronic low back pain and suspected or diagnosed axSpA and was designed as a post hoc analysis: the Sacroiliac joint MAgnetic resonance imaging and Computed Tomography study with 110 patients²² and the Virtual Non-Calcium Susceptibility-Weighted Imaging study with 72 patients.²³ For the present analysis, patients were divided into two groups based on the clinical diagnosis established by the local expert rheumatologists considering all imaging findings in conjunction with clinical and laboratory results—a group of patients with a diagnosis of axSpA and a control group of patients with other diagnoses (eg, degenerative or mechanical back pain or non-specific low back pain). Patients or individuals with no complete imaging or clinical data were excluded from the analysis (figure 1).

Readers and scoring system

XR, MRI and CT of the SIJs were anonymised prior to the reading sessions. In a virtual session, all readers were introduced to the reading methods and scoring system before starting the reading process. In a first reading session, nine readers scored XR, MRI (oblique-coronal T1-weighted and short-tau inversion recovery (STIR) sequence) and CT separately for the presence or absence of axSpA using a dichotomous score ('axSpA' or 'non-axSpA'). In a separate reading session, the readers scored combined XR&MRI and CT&MRI. The readers had different years of experience in musculoskeletal imaging and were divided into three reader groups based on their experience level: an inexperienced reader group (low-XP) including three medical research students (CS, FR and DD) with 0 to 1 year of experience, a semi-experienced reader group (intermediate-XP) consisting

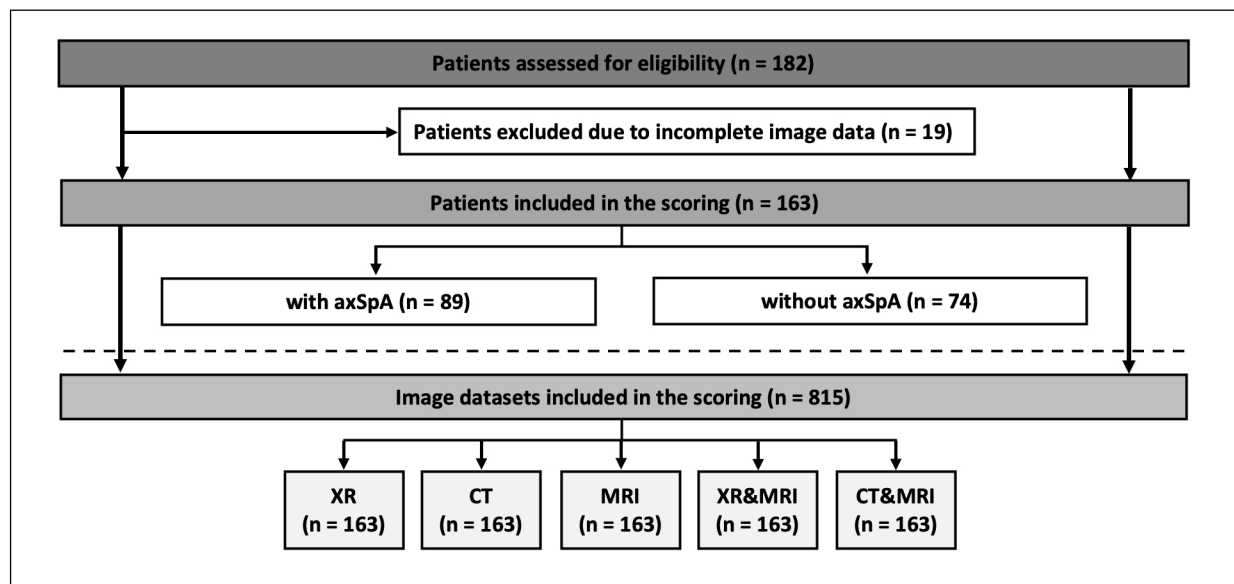


Figure 1 Flow chart of study inclusion. axSpA, axial spondyloarthritis; XR, X-ray.

of radiologists with 3–8 years of experience (STU, JG and KZ), and an experienced group of senior physicians (two radiologists (TD and IE) and one senior rheumatologist (DP)) with 12–17 years of experience in musculoskeletal imaging (high-XP). The readers were blinded to clinical information and other imaging data.

Statistical analysis

Statistical analysis was performed using GraphPad Prism (V.9.4.1 for MacOS, GraphPad Software, La Jolla, California, USA) and Microsoft Excel (V.16.65 for MacOS). Parameters of diagnostic accuracy for the different imaging modalities (XR, CT and MR) and their combinations (XR&MR and CT&MR) were calculated using contingency tables, separately for the three reader groups: sensitivity, specificity, positive/negative predictive values and likelihood ratios (LR+/LR-) using the

Wilson/Brown method. The McNemar test was used to assess significant differences in correct and incorrect answers between the reader groups separately for each imaging modality and combination. Furthermore, Fleiss kappa (κ) was calculated to quantify the agreement of the three readers within each of the reader groups. Fleiss κ was interpreted according to Landis and Koch.²⁴ A p value <0.05 was considered statistically significant.

RESULTS

Subjects

A total of 182 subjects underwent eligibility assessment, with 19 patients excluded due to insufficient clinical or imaging data. After applying the exclusion criteria, 163 subjects with low back pain (82 women; mean age, 38 (SD 10.6), 19–62 years and a mean duration of low back pain

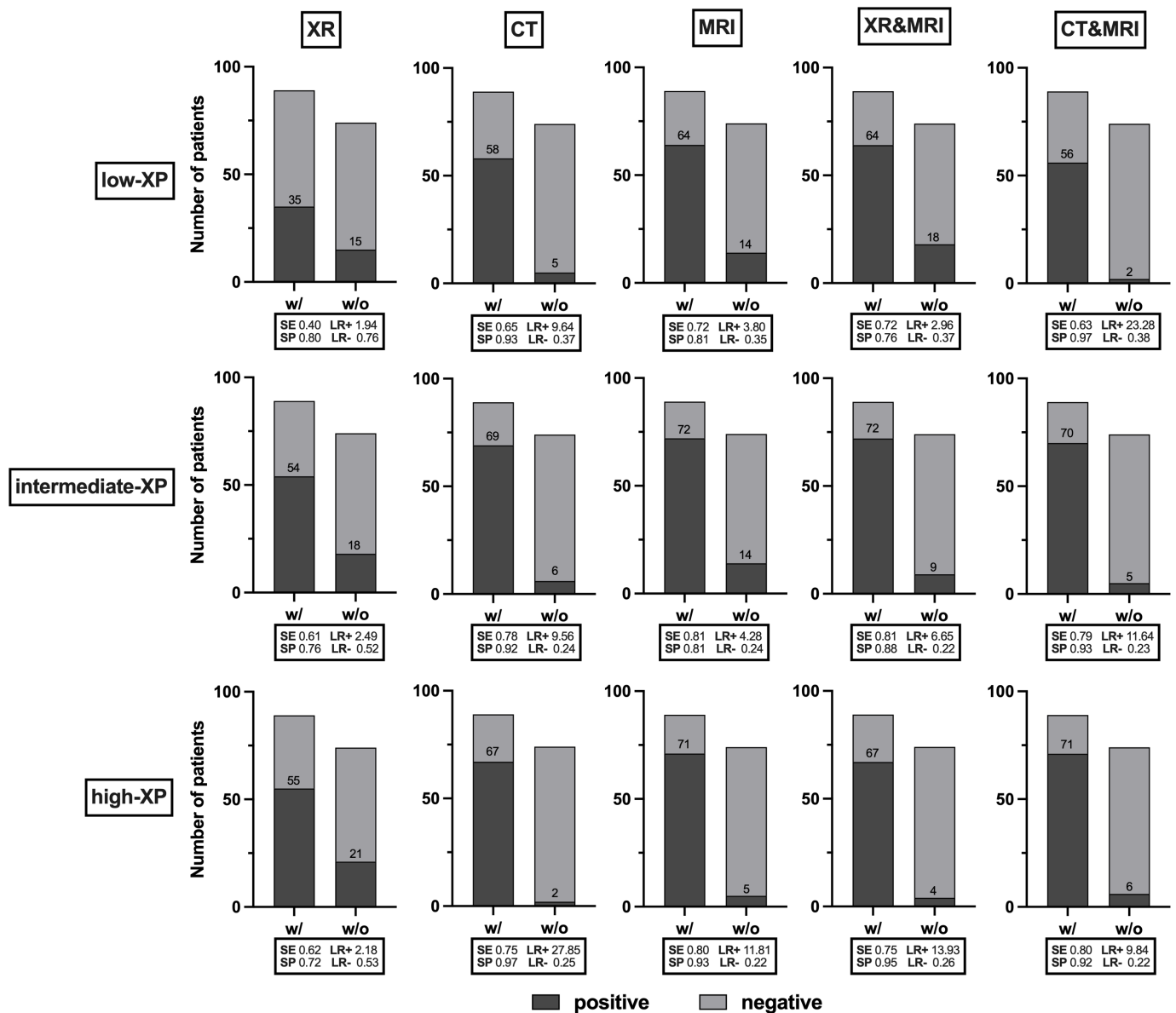


Figure 2 Results of the contingency table analyses. high-XP, experienced reader group; intermediate-XP, semi-experienced reader group; low-XP, inexperienced reader group; LR-, negative likelihood ratio; LR+, positive likelihood ratio; SE, sensitivity; SP, specificity; w/, with axial spondyloarthritis; w/o, without axial spondyloarthritis; XR, X-ray.

of 80.6 months (SD 89.9)) were included in the analysis. Prior to the reading sessions, 89 patients were diagnosed with axSpA by the local rheumatologist, while 74 patients were diagnosed with other conditions (56 patients with degenerative or mechanical SIJ disease and 18 patients with non-specific low back pain) (figure 2). The mean Bath Ankylosing Spondylitis Disease Activity Index was 4.6 (SD 1.8) in patients with axSpA with a mean C reactive protein level of 7.8 mg/dL (SD 12.0). HLA-B27 was positive in 60.2% of the patients. Further clinical details are available in our earlier publication.¹³

Image reading

Overall, 815 imaging datasets were anonymised before scoring. Although diagnostic accuracy was higher for high-XP compared with the intermediate-XP group, except for XR, the difference was not statistically significant. Diagnostic accuracies of the intermediate-XP and high-XP group were also not significantly different in the combined evaluation of XR&MR (84.0%) and CT&MR (85.3%). In the low-XP and intermediate-XP groups, the highest diagnostic accuracy was found for CT&MR (78.5% and 85.3%, respectively), while the high-XP readers achieved the highest accuracy for MRI (85.9%). Figure 2 summarises the results of the contingency table analyses. The contingency tables are presented in online supplemental table 1. Diagnostic accuracies differed most significantly for MRI between low-XP and high-XP ($p=0.001$), followed by XR&MR ($p=0.009$) (figure 3). Examples of the image reading across the different reader groups are shown in figure 4. Regarding the inter-rater reliability, we found higher kappa values for CT in all three reader

groups compared with the other modalities and combinations. Agreement between the readers was best for CT in high-XP with $\kappa=0.87$, followed by $\kappa=0.79$ for intermediate-XP. The low-XP readers achieved only fair agreement, except for CT ($\kappa=0.50$) and CT&MR ($\kappa=0.41$).

DISCUSSION

In this study, we investigated differences in diagnostic accuracies achieved in the interpretation of musculoskeletal imaging findings typical of axSpA in the SIJs by three groups of readers with different levels of experience. Our results show that the level of experience and the imaging modality (XR, CT, MRI) has an impact on the diagnostic accuracy achieved by a reader. As expected, an experienced reader had the highest diagnostic accuracy in all modalities, except for XR. The overall highest accuracy across all three reader groups was achieved for CT while the accuracy of MRI interpretation varied most markedly with the experience level. Furthermore, our study highlights that the addition of CT to MRI resulted in the highest diagnostic accuracy in both the inexperienced and the semi-experienced reader group. Interestingly, this combined modality approach had no additive value for the three highly experienced readers, suggesting that the benefits of adding CT are more pronounced in readers with less experience. This finding underscores the potential of complementary imaging modalities to enhance diagnostic accuracy, particularly in less-specialised centres. These insights contribute valuable considerations for optimising the diagnostic approach in the assessment of axSpA.

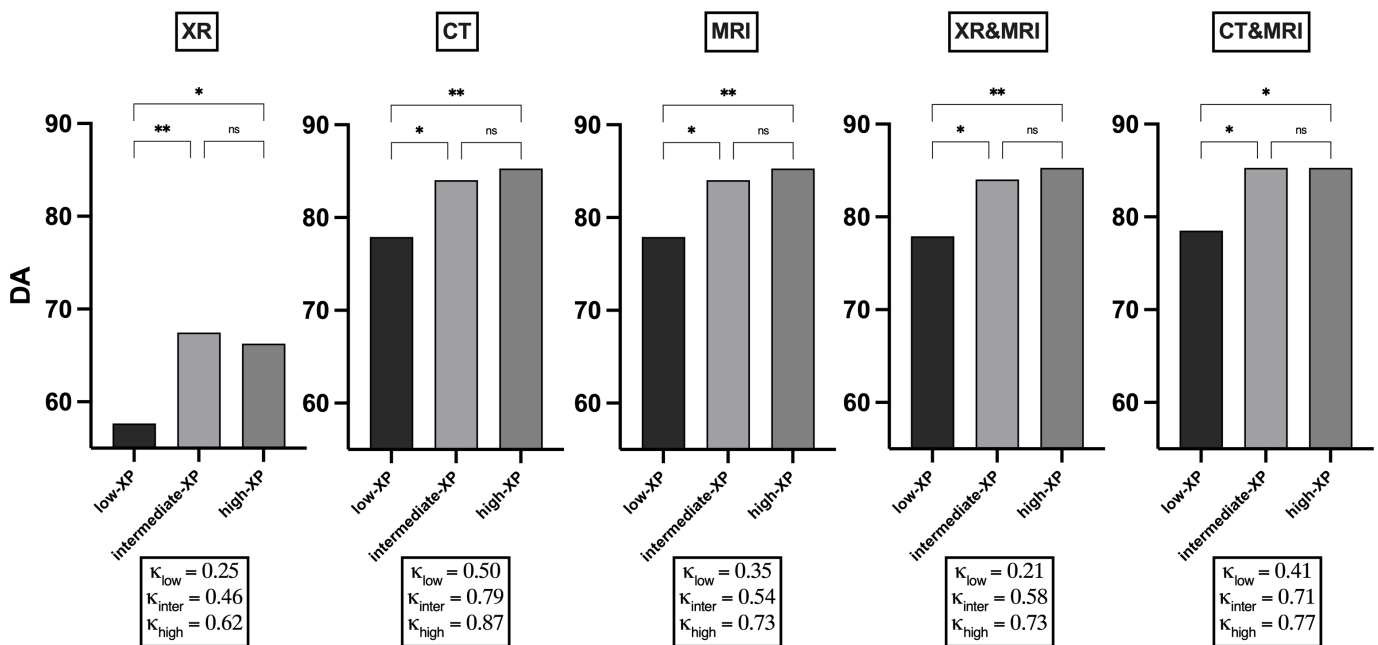


Figure 3 Box plots of diagnostic accuracies (DA) by modality and level of reader experience. high-XP, experienced reader group; intermediate-XP, semi-experienced reader group; κ , Fleiss kappa, low-XP, inexperienced reader group; XR, X-ray. Not significant (ns): $p>0.05$, * $p\leq 0.05$, ** $p\leq 0.01$. The greatest difference was found for MRI between low-XP and high-XP ($p=0.001$), followed by XR&MR ($p=0.009$). Almost perfect agreement was shown for CT in high-XP with $\kappa=0.87$. The best agreement in intermediate-XP was also found for CT with $\kappa=0.79$.

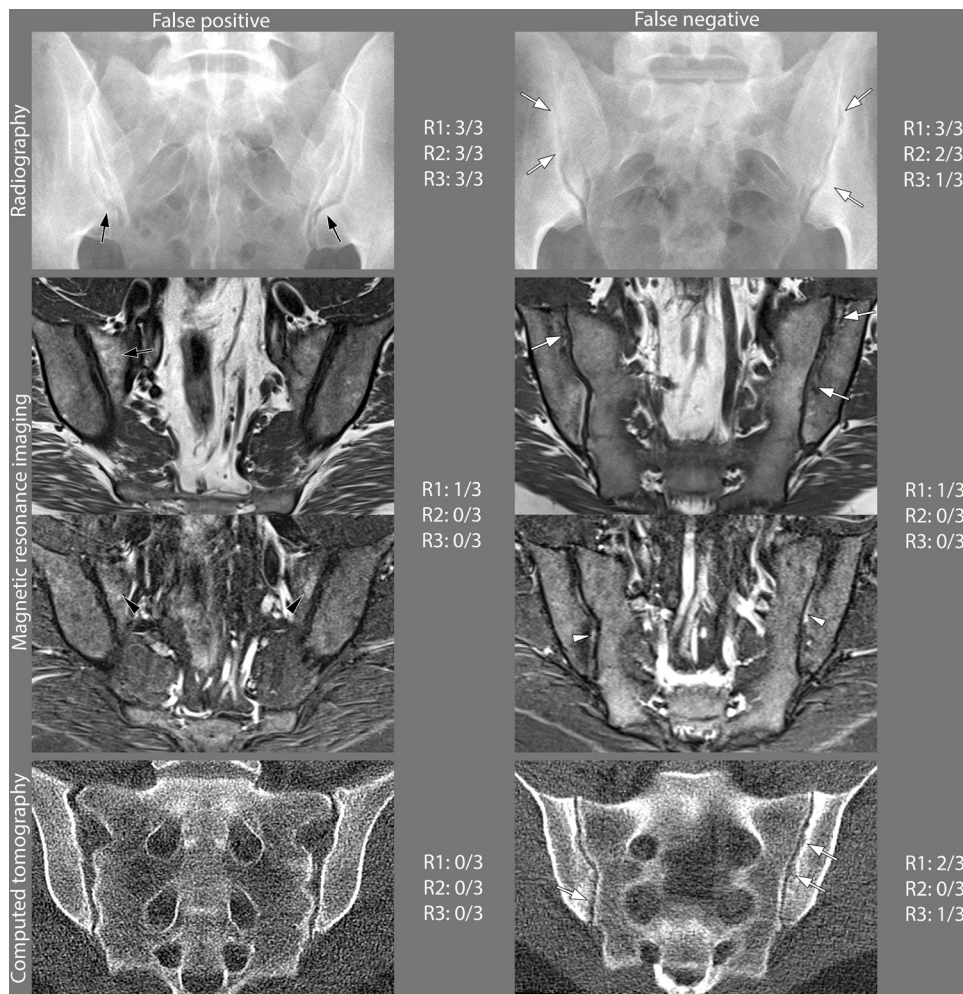


Figure 4 Examples of the image reading across different reader groups. Image examples of a patient with false positive scoring for axial spondyloarthritis are presented in the left column. In the right column, imaging examples of a patient with false negative scoring for axial spondyloarthritis are presented. R1=inexperienced reader group, R2=semi-experienced reader group, R3=experienced reader group. In radiography, all readers misinterpreted the mild sclerosis and discreet erosive changes in the ventral portion of the sacroiliac joints as axial spondyloarthritis (black arrows). In comparison with radiography, the majority of readers were able to accurately identify the structural changes in a patient with axial spondyloarthritis in MRI and CT (white arrows).

Tailoring the appropriate imaging modality is essential for diagnosis in axSpA. Therefore, knowledge about the strengths and weaknesses of the different imaging modality is essential.²⁰ In clinical practice, the assessability and reliability of radiological findings are crucial for selecting the best modality in a given situation. Special training and experience are required, which is why the diagnosis is often delayed outside of specialised axSpA centres. This underlines the importance of continuous training of young and inexperienced radiologists in the field of axSpA imaging using established and new methods.^{25–27} Our results make a crucial contribution to overcoming existing shortcomings, as our findings reveal that, besides the experience of the reader itself, it is the choice of modality that has an important impact on overall diagnostic accuracy. However, the diagnosis cannot be established by imaging alone but clinical features need also consideration in the diagnostic

process, as even highly experienced readers only achieve an accuracy of 85.9%.

Since its introduction in axSpA imaging, MRI has gained importance for early diagnosis and is currently being discussed as the first-line imaging modality whenever axSpA is suspected. However, our results suggest that accurate interpretation of MRI heavily depends on the reader's experience and requires high-level training compared with CT. Furthermore, MRI is far more expensive than CT and XR and is also less widely available, which may have an overall impact on its usability, particularly in outpatient practices among radiologists without specific experience in axSpA imaging. CT seems to be easier to interpret, especially for beginners, resulting in similar diagnostic accuracy but better inter-rater reliability, which makes it more suitable for inexperienced readers. The disadvantage of radiation exposure in CT is eliminated by the advent of low-dose CT protocols, which

expose the patient to radiation doses comparable to XR. A marginal increase in diagnostic accuracy was shown for XR from low-XP to intermediate-XP and high-XP group. Still, the performance and reliability of XR are inferior to those of the cross-sectional imaging techniques across all experience levels, further underlining that XR is difficult to interpret and is especially unsuited for inexperienced readers. Exercises to improve the interpretation of XR images of the SIJs have also been carried out in the past and showed poor agreement even among recognised experts in the disease.²⁸ With XR, there is a risk of misinterpreting imaging findings and missing subtle alterations. This has not been improved over the decades, even through a broader understanding of cross-sectional imaging techniques.

Our results show MRI to be the imaging modality with the best learnability. Hence, training and experience can lead to a significant improvement in diagnostic accuracy when MRI is used²⁹ while, for XR, the impact of training seems to be limited. Indeed, previous studies have reported its limited reproducibility in axSpA imaging,³⁰ showing large interobserver and intraobserver variations.³¹ Several studies report advantages of CT and MRI over XR in axSpA diagnosis.³² However, inexperienced readers show a more constantly good performance for the detection of typical axSpA-related abnormalities on CT, which makes it interesting as an imaging modality in non-specialized centres.

One limitation of our study is that we only used SIJ images in the reading sessions. Therefore, the effect on diagnostic accuracy of using spine imaging remains to be investigated. Furthermore, we did not separately investigate the influence of known clinical data on diagnostic accuracy. The reader sample included one rheumatologist with extensive scientific experience in axSpA imaging but no specific training in radiology. As the inter-rater agreement was high, we believe that experience is more important than profession and that this fact has not a major impact on our data. Specifically, a comparison of rheumatologists and radiologists was not intended in this study. While this method poses a potential bias risk due to circular reasoning, considering that some of the readers might have participated in the diagnostic process within the clinical setting, it is crucial to note that the reading outcomes from this study were distinct from the clinical diagnostic process and thorough anonymisation was implemented. In the absence of a widely accepted reference standard in axSpA studies, an expert consensus could potentially establish a reference for imaging positivity. However, in this context, the alignment with the final clinical outcome was deemed more important than with imaging criteria alone. Our assessment also does not allow for a differentiated evaluation of the experience-dependent influence of individual inflammatory or structural lesions on the image evaluations. However, it seems reasonable that the interpretation of the large number of detectable, especially inflammatory, lesions in MRI needs more experience than the simple structural assessment

with CT or XR. Lastly, measuring experience in axSpA imaging accurately is difficult, as years of experience may not correlate perfectly with the number of scans read.

In conclusion, our findings suggest that the diagnostic accuracy in interpreting images of the SIJs in patients with suspected axSpA varies based on the experience level of radiologists and the chosen imaging modality. It is noteworthy that, overall, the impact of experience on diagnostic accuracy is most pronounced in MRI. However, it is observed that even inexperienced radiologists with a short duration of training demonstrate good diagnostic accuracy when interpreting CT. Moreover, the diagnosis of axSpA relies on both clinical and imaging in the diagnostic process.

Author affiliations

¹Department of Radiology, Charité Universitätsmedizin Berlin Campus Charité Mitte, Berlin, Germany

²Department of Hematology, Oncology, Tumor Immunology, Charité Universitätsmedizin Berlin Campus Charité Mitte, Berlin, Germany

³Diagnostic Imaging, Sheba Medical Center, Tel Hashomer, Israel

⁴Tel Aviv Sourasky Medical Center, Tel Aviv, Israel

⁵Department of Rheumatology, Charité Universitätsmedizin Berlin Campus Charité Mitte, Berlin, Germany

⁶Department of Gastroenterology, Infectiology and Rheumatology (including Nutrition Medicine), Charité Universitätsmedizin Berlin Campus Benjamin Franklin, Berlin, Germany

⁷Berlin Institute of Health at Charité, Berlin, Germany

Twitter Robert Biesen @BiesenRobert, Mikhail Protopopov @mprotopopov and Fabian Proft @ProftDr

Acknowledgements The authors thank Ms Bettina Herwig for language editing. The authors thank the Berlin Institute of Health for personal funding (STU, JR and TD) and providing essential infrastructure for data collection. This research project was supported by the Assessment of Spondyloarthritis international Society (ASAS) (research grant for KZ).

Contributors FR: design of scoring system, image scoring, data evaluation, article draft, critical revision of the manuscript for important intellectual content. KZ, IE, JG, DD, CS: image scoring, critical revision of the manuscript for important intellectual content. RB: patient acquisition, data collection, critical revision of the manuscript for important intellectual content. HH: patient acquisition, data collection, critical revision of the manuscript for important intellectual content. VR: patient acquisition, data collection, critical revision of the manuscript for important intellectual content. JR: patient acquisition, data collection, critical revision of the manuscript for important intellectual content. MP: patient acquisition, data collection, critical revision of the manuscript for important intellectual content. FP, K-GAH: patient acquisition, data collection, critical revision of the manuscript for important intellectual content. DP: patient acquisition, data evaluation, critical revision of the manuscript for important intellectual content. TD: image scoring, data evaluation, critical revision of the manuscript for important intellectual content. STU: conception and design of the study, design of scoring system, image scoring, data evaluation, statistical calculations, article draft, critical revision of the manuscript for important intellectual content. STU acts as the guarantor of the study.

Funding This research project was funded by the Assessment of Spondyloarthritis international Society (ASAS) and the Berlin Institute of Health. The funding sources were not involved in study design, in the collection, analysis and interpretation of data, in the writing of the report or in the decision to submit the paper for publication.

Competing interests KZ reports funding (research grant) from the Assessment of Spondyloarthritis international Society (ASAS) during the conduct of this study. IE reports personal fees from AbbVie, Eli Lilly and Novartis. RB reports personal fees from AstraZeneca, Galapagos, GlaxoSmithKline, Medac and Novartis. HH reports grants from Sobi and personal fees from AbbVie, Novartis, Pfizer, Roche and UCB outside the submitted work. JR is participant in the BIH-Charité Clinician Scientist Programme funded by the Charité—Universitätsmedizin Berlin and the Berlin Institute of Health. FP reports grants and personal fees from Novartis, Lilly and UCB, as well as personal fees from AbbVie, Amgen, BMS, Hexal, Janssen, MSD, Pfizer

and Roche. K-GAH reports personal fees from AbbVie, MSD, Pfizer and Novartis, he is also the co-founder of BerlinFlame. DP reports grants and personal fees from AbbVie, Eli Lilly, MSD, Novartis and Pfizer and personal fees from Bristol-Myers Squibb, Roche, UCB, Biocad, GlaxoSmithKline and Gilead outside the submitted work. TD reports personal fees from MSD, Novartis and Eli Lilly and reports funding from the Berlin Institute of Health (BIH) during the conduct of this study. STU reports funding from BIH during the conduct of this study (Junior Digital Clinician Scientist Program). All other authors report no funding.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval The institutional ethics review board approved all investigations prior to study commencement (EA1/0886/16; EA1/073/10). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Katharina Ziegeler <http://orcid.org/0000-0002-9763-2420>
 Iris Eshed <http://orcid.org/0000-0002-4655-9606>
 Robert Biesen <http://orcid.org/0000-0002-0434-7832>
 Judith Rademacher <http://orcid.org/0000-0001-7442-2570>
 Mikhail Protopopov <http://orcid.org/0000-0003-4840-5069>
 Fabian Proft <http://orcid.org/0000-0003-4306-033X>
 Kay Geert A Hermann <http://orcid.org/0000-0001-6142-3814>
 Denis Poddubnyy <http://orcid.org/0000-0002-4537-6015>
 Torsten Diekhoff <http://orcid.org/0000-0003-3593-1449>
 Sevtap Tugce Ulas <http://orcid.org/0000-0003-3871-3746>

REFERENCES

- Maksymowych WP, Lambert RG, Østergaard M, *et al.* MRI lesions in the sacroiliac joints of patients with spondyloarthritis: an update of definitions and validation by the ASAS MRI working group. *Ann Rheum Dis* 2019;78:1550–8.
- Sieper J, Poddubnyy D. Axial spondyloarthritis. *Lancet* 2017;390:73–84.
- Braun J, Bollow M, Remlinger G, *et al.* Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. *Arthritis Rheum* 1998;41:58–67.
- Sieper J, Rudwaleit M, Khan MA, *et al.* Concepts and epidemiology of spondyloarthritis. *Best Pract Res Clin Rheumatol* 2006;20:401–17.
- Sieper J, Rudwaleit M, Baraliakos X, *et al.* The assessment of spondyloarthritis international society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 2009;68 Suppl 2:ii1–44.
- Navarro-Compán V, Sepriano A, El-Zorkany B, *et al.* Axial spondyloarthritis. *Ann Rheum Dis* 2021;80:1511–21.
- Vleeming A, Schuenke MD, Masi AT, *et al.* The sacroiliac joint: an overview of its anatomy, function and potential clinical implications. *J Anat* 2012;221:537–67.
- Baraliakos X. Imaging in axial spondyloarthritis. *Isr Med Assoc J* 2017;19:712–8.
- Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing Spondylitis: do we need new criteria. *Arthritis Rheum* 2005;52:1000–8.
- van der Heijde D, Rudwaleit M, Landewé RBM, *et al.* Justification for including MRI as a tool in the diagnosis of axial SpA. *Nat Rev Rheumatol* 2010;6:670–2.
- Mau W, Zeidler H, Mau R, *et al.* Clinical features and prognosis of patients with possible ankylosing spondylitis. results of a 10-year followup. *J Rheumatol* 1988;15:1109–14.
- Diekhoff T, Lambert R, Hermann KG. “MRI in axial spondyloarthritis: understanding an ‘ASAS-positive MRI’ and the ASAS classification criteria”. *Skeletal Radiol* 2022;51:1721–30.
- Diekhoff T, Eshed I, Radny F, *et al.* Choose wisely: imaging for diagnosis of axial spondyloarthritis. *Ann Rheum Dis* 2022;81:237–42.
- Rudwaleit M, Jurik AG, Hermann K-GA, *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;68:1520–7.
- Eshed I, Miloh-Raz H, Dulitzki M, *et al.* Peripartum changes of the sacroiliac joints on MRI: increasing mechanical load correlating with signs of edema and inflammation kindling spondyloarthropathy in the genetically prone. *Clin Rheumatol* 2015;34:1419–26.
- de Koning A, de Bruin F, van den Berg R, *et al.* Low-dose CT detects more progression of bone formation in comparison to conventional radiography in patients with ankylosing spondylitis: results from the SIAS cohort. *Ann Rheum Dis* 2018;77:293–9.
- de Bruin F, de Koning A, van den Berg R, *et al.* Development of the CT syndesmophyte score (CTSS) in patients with ankylosing spondylitis: data from the SIAS cohort. *Ann Rheum Dis* 2018;77:371–7.
- Lambert RGW, Hermann KGA, Diekhoff T. Low-dose computed tomography for axial spondyloarthritis: update on use and limitations. *Curr Opin Rheumatol* 2021;33:326–32.
- Diekhoff T, Hermann KGA, Lambert RG. Future of low-dose computed tomography and dual-energy computed tomography in axial spondyloarthritis. *Curr Rheumatol Rep* 2022;24:198–205.
- van Gaalen FA, Bakker PAC, de Hooge M, *et al.* Assessment of sacroiliitis by radiographs and MRI: where are we now. *Curr Opin Rheumatol* 2014;26:384–8.
- Poddubnyy D, Diekhoff T, Baraliakos X, *et al.* Diagnostic evaluation of the sacroiliac joints for axial spondyloarthritis: should MRI replace radiography. *Ann Rheum Dis* 2022;81:1486–90.
- Diekhoff T, Greese J, Sieper J, *et al.* Improved detection of erosions in the sacroiliac joints on MRI with volumetric interpolated breath-hold examination (VIBE): results from the SIMACT study. *Ann Rheum Dis* 2018;77:1585–9.
- Deppe D, Hermann K-G, Proft F, *et al.* CT-like images of the sacroiliac joint generated from MRI using susceptibility-weighted imaging (SWI) in patients with axial spondyloarthritis. *RMD Open* 2021;7:e001656.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159.
- Muşetescu AE, Bobircă A, Gherghina FL, *et al.* Interobserver reliability of magnetic resonance imaging of Sacroiliac joints in axial spondyloarthritis. *Life* 2022;12:470.
- Hermann KG, Protopopov M, Serfaty A, *et al.* Pos1460 contributing to the training of imaging in rheumatology by experts worldwide via interactive mobile E-teaching: berlintcaseviewer. *Ann Rheum Dis* 2022;81(Suppl 1):1075.
- Hadsbjerg AEF, Østergaard M, Paschke J, *et al.* Pos0918 validation of a modified New York criteria retic E-tool for increasing scoring proficiency of radiographs of the sacroiliac joints in axial spondyloarthritis. *ARD*; 2023:82 769–70
- 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;62:519–25.
- Garcia Salinas R, Marin J, Aguilar G, *et al.* Recognition of spondyloarthritis lesions on magnetic resonance imaging: results of an educational intervention. *Int J Rheum Dis* 2020;23:1651–5.
- Christiansen AA, Hendricks O, Kuettel D, *et al.* Limited reliability of radiographic assessment of sacroiliac joints in patients with suspected early spondyloarthritis. *J Rheumatol* 2017;44:70–7.
- Yazici H, Turunç M, Ozdoğan H, *et al.* Observer variation in grading sacroiliac radiographs might be a cause of ‘sacroiliitis’ reported in certain disease states. *Ann Rheum Dis* 1987;46:139–45.
- Battafarano DF, West SG, Rak KM, *et al.* Comparison of bone scan, computed tomography, and magnetic resonance imaging in the diagnosis of active sacroiliitis. *Semin Arthritis Rheum* 1993;23:161–76.