DISSERTATION

Der Einfluss der klinischen Information auf die radiologische Befundung bei Patient*innen mit Verdacht auf axiale Spondyloarthritis

Axial spondyloarthritis: The role of clinical information in imaging evaluation.

zur Erlangung des akademischen Grades

Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät

Charité – Universitätsmedizin Berlin

von

Tim Alexander Pohlner

Erstbetreuung: Herr Professor Dr. Denis Poddubnyy

Datum der Promotion: 20.06.2025

Vorwort:

Teileergebnisse dieser Arbeit wurden vorab bereits in einer Fachzeitschrift (RMD Open) veröffentlicht worden.

Publikation:

Pohlner T, Deppe D, Ziegeler K, Proft F, Protopopov M, Rademacher J, Rios Rodriguez V, Torgutalp M, Braun J, Diekhoff T, Poddubnyy D. Diagnostic accuracy in axial spondyloarthritis: a systematic evaluation of the role of clinical information in the interpretation of sacroiliac joint imaging. RMD Open. 2024;10(2).

Inhaltsverzeichnis

Vorwort: I

Abbildungsverzeichnis V

Tabellenverzeichnis VI

Abkürzungsverzeichnis VII

Zusammenfassung IX

Abstract X

1 Einleitung 1

1.1 Axiale Spondyloarthritis 1

1.1.1 Überblick 1

1.1.2 Epidemiologie 2

1.1.3 Pathogenese 3

1.1.4 Klinisches Bild 3

1.2 Diagnostik und Klassifikation der axSpA 6

1.2.1 Klinische Diagnostik 6

1.2.2 Anamnese 6

1.2.3 Körperliche Untersuchung und Labordiagnostik 7

1.2.4 Bildgebung 7

1.2.5 Konventionelle Röntgenuntersuchung 8

1.2.6 Magnetresonanztomografie (MRT) 9

1.2.7 ASAS-Gruppe 10

1.2.8 Die ASAS-Klassifikationskriterien für axSpA 11

1.2.9 Sakroiliitis in der MRT: ASAS-Definitionen 12

1.2.10 Berlin Algorithmus 13

1.2.11 Diagnoseverzögerung 14

1.3 Therapie und Verlauf 14

1.4 Die Rolle der klinischen Information in der radiologischen Beurteilung 14

1.5 Fragestellung 15
2 Material und Methoden 17
2.1 Material 17
2.1.1 Auswahl des Patient*innenkollektivs 17
2.1.2 Radiolog*innen 19
2.2 Radiologische Befundung 21
2.3 Statistische Auswertung 23
3 Ergebnisse 27
3.1 Demografische Daten und klinische Befunde 27
3.2 Ergebnisse der radiologischen Befundung 28
3.3 Auswertung der radiologischen Befundung: Die diagnostische Genauigkeit 30
3.4 Befundung der Röntgenbilder nach den mNYC 35
3.5 Befundung der MRT-Aufnahmen: SpA-typische Veränderungen 36
3.6 Die subjektiv geschätzte diagnostische Sicherheit 39
3.7 Der subjektiv geschätzte Einfluss der klinischen Information auf die Befundung 44
3.8 Analyse der Befundung in Abhängigkeit von der subjektiv geschätzten diagnostischen Sicherheit 46
4 Diskussion: 53
4.1 Diagnostische Sicherheit und Einfluss der klinischen Information: 55
4.2 Variabilität der Reader: 56
4.3 Argumente pro und contra klinische Information 57
4.4 Stellenwert der bildgebenden Methoden bei axSpA 59
4.5 Was könnten diese Ergebnisse für die klinische Praxis bedeuten? 61
5 Limitationen 62
Quellenverzeichnis 63
Appendix 69
Eidesstattliche Versicherung 73

Anteilserklärung der erfolgten Publikation 74

Lebenslauf 75

Publikationsliste 76

Danksagung 77

Abbildungsverzeichnis

Abbildung 1 ASAS Klassifikationskriterien 12
Abbildung 2 Übersicht: Selektion der Studienpopulation 20
Abbildung 3 Übersicht über die vier Runden der radiologischen Auswertung. 21
Abbildung 4 Veränderungen der Befundung durch die klinische Information 32
Abbildung 5 Strukturelle Röntgenveränderungen der SIG 35
Abbildung 6 axSpA typische SIG-Veränderungen in der MRT 37
Abbildung 7 Diagnostische Sicherheit: Röntgen 39
Abbildung 8 Diagnostische Sicherheit: Röntgen- und MRT-Auswertung 41
Abbildung 9 Einfluss der klinischen Information Röntgen vs. Röntgen und MRT 44
Abbildung 10 Diagnostische Sicherheit und Einfluss der klinischen Information pro Reader (Röntgen) 45
Abbildung 11 Diagnostische Sicherheit und Einfluss der klinischen Information pro Reader (Röntgen und MRT) 46
Abbildung 12 Diagnostische Genauigkeit in Abhängigkeit der Sicherheit: Röntgen 47
Abbildung 13 Diagnostische Genauigkeit in Abhängigkeit der Sicherheit: Röntgen und MRT 49

Appendix-Abbildungen:

Abbildung A 1 Befundmaske der Runde 1 69
Abbildung A 2 Liste der klinischen Informationen 70
Abbildung A 3: ASAS modifizierter Berlin Algorithmus 71
Abbildung A 4 ASAS-Empfehlungen: Radiologischer Befund: 72

Tabellenverzeichnis

Tabelle 1 Charakteristika des entzündlichen Rückenschmerzes. 4

Tabelle 2 Graduierung der strukturellen SIG-Veränderungen nach den modifizierten New York Kriterien 9

Tabelle 3 Demografische und klinische Charakteristika der Patient*innen 28

Tabelle 4 Häufigkeiten SpA-typischer radiologischer Veränderungen 29

Tabelle 5 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen-Befundung 30

Tabelle 6 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen- und MRT-Befundung. 31

Tabelle 7 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen-Befundung (separat pro Reader) 33

Tabelle 8 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen- und MRT-Befundung (separat pro Reader) 34

Tabelle 9 Aktive Entzündung und strukturelle Läsionen (MRT-Befundung) 38

Tabelle 10 Diagn. Sicherheit: Röntgen ohne klinische Information (separat pro Reader) 40

Tabelle 11 Diagn. Sicherheit: Röntgen mit klinischer Information (separat pro Reader) 40

Tabelle 12 Diagn. Sicherheit Röntgen und MRT ohne klin. Information (separat pro Reader) ... 42

Tabelle 13 Diagn. Sicherheit Röntgen und MRT mit klin. Information (separat pro Reader) 42

Tabelle 14 Änderungen der diagn. Sicherheit durch klinische Information (nur Röntgen und Röntgen + MRT) 43

Tabelle 15 Änderungen der diagn. Sicherheit durch klin. Information (separat pro Reader) 43

Tabelle 16 Röntgen-Befundung in Abhängigkeit der subjektiv geschätzten diagn. Sicherheit 48

Tabelle 17 Röntgen- und MRT-Befundung in Abhängigkeit der subjektiv geschätzten diagn. Sicherheit 50

Tabelle 18 Röntgen-Befundung in Abhängigkeit der subjektiv geschätzten diagnostischen Sicherheit separat pro Reader 51

Tabelle 19 Röntgen- und MRT-Befundung in Abhängigkeit der subjektiv geschätzten diagnostischen Sicherheit separat pro Reader 52

Abkürzungsverzeichnis

CED	chronisch entzündliche Darmerkrankungen
HLA	humanes Leukozytenantigen
SpA	Spondyloarthritis
axSpA	axiale Spondyloarthritis
r-axSpA	röntgenologische axiale Spondyloarthritis
nr-axSpA	nicht-röntgenologische axiale Spondyloarthritis
AS	ankylosierende Spondylitis (Spondylitis ankylosans)
BME	bone marrow edema
KMÖ	Knochenmarködem
ASAS	Assesment of SpondyloArthritis international Society
ESSR	European Society of Musculoskeletal Radiology
EULAR	European League Against Rheumatism
CT	Computertomografie
MRT	Magnetresonanztomografie
mNYC	modified New York Criteria (modifizierte New York Kriterien)
STIR	Short Tau Inversion Recovery
SIG	Sakroiliakalgelenk
MHC	Major Histocompatibility Complex (Haupthistokompatibilitätskomplex)
ERAP	Endoplasmatisches Retikulum Aminopeptidase
IL	Interleukin
TNF	Tumornekrosefaktor
JAK	Januskinase
PCR	Polymerasekettenreaktion
FACS	Fluorescence Activated Cell Sorting (Durchflusszytometrie)
CRP	C-reaktives Protein
BSG	Blutsenkungsgeschwindigkeit
WS	Wirbelsäule
ASDAS	Ankylosing Spondylitis Disease Activity Score

BASDAI	Bath Ankylosing Spondylitis Disease Activity Index
csDMARD	Conventional synthetic disease modifying antirheumatic drug
PPV	Positiver prädiktiver Wert
NPV	Negativer prädiktiver Wert
klin.	klinisch
diagn.	diagnostisch

Zusammenfassung

Ziel: Röntgen- und Magnetresonanztomographie (MRT)-Aufnahmen der Sakroiliakalgelenke (SIG) sind für die Diagnose und Klassifizierung von Patient*innen mit axialer Spondyloarthritis (axSpA) unerlässlich. Ziel dieser Studie war es, den Einfluss klinischer Informationen auf die Genauigkeit der Bildinterpretation zu untersuchen.

Methoden: Die Röntgen- und MRT-Bilder von 109 Patient*innen mit Verdacht auf axSpA wurden separat durch drei Radiolog*innen in vier aufeinanderfolgenden Runden ausgewertet. Dabei wurden zunächst nur die Röntgenaufnahmen, später Röntgen- und MRT-Aufnahmen jeweils ohne und mit klinischer Information (demografische Daten, SpA-Merkmale, körperliche Aktivität, Schwangerschaft) über die Patient*innen analysiert. Die Befundung der Röntgenbilder erfolgte anhand der modifizierten New York Kriterien, die der MRT-Aufnahmen nach entzündlichen und/oder strukturellen Veränderungen der SIG sowohl separat pro Radiolog*in als auch kondensiert nach dem Mehrheitsprinzip (≥2/3 Radiolog*innen). Es wurde die Kompatibilität (Präzision) der radiologischen Befunde mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard) sowie Sensitivität, Spezifität, PPV, NPV und die Übereinstimmung der Radiolog*innen untersucht. Außerdem wurden die selbst geschätzte diagnostische Sicherheit der Radiolog*innen und der geschätzte Einfluss der klinischen Information auf die Befundung erhoben und in Abhängigkeit von der Bildmodalität und klinischen Information ausgewertet.

Ergebnisse: Die Präzision der Röntgen-Befunde ohne bzw. mit klinischer Information stieg von 70 % auf 78 % und für Röntgen und MRT kombiniert von 81 % auf 82 %. Bei alleiniger Befundung der Röntgenaufnahmen betrugen Sensitivität und Spezifität 51 % und 94 % ohne und 60 % und 100 % mit klinischer Information, während für die gemeinsame Beurteilung von Röntgen- und MRT-Aufnahmen 74 % und 90 % bzw. 71 % und 98 % erzielt wurden. Die selbst geschätzte Diagnosesicherheit der Radiolog*innen stieg von 5.2 ± 1.9 auf 6.0 ± 1.7 mit klinischer Information für die Röntgenbefunde bzw. von 6.7 ± 1.6 auf 7.2 ± 1.6 für die Röntgen- und MRT-Auswertung. Die selbst geschätzte diagnostische Sicherheit stieg bei allen Bildmodalitäten durch die klinische Information an, während der Einfluss der klinischen Information auf Röntgen- bzw. Röntgen- und MRT-Befundung kombiniert nicht unterschiedlich eingeschätzt wurde.

Schlussfolgerung: Die Präzision, die Spezifität und die diagnostische Sicherheit der radiologischen Auswertung nahm zu, wenn klinische Informationen zur Verfügung gestellt wurden. Diese Daten sprechen für die Bereitstellung klinischer Informationen für Radiolog*innen.

Abstract

Objectives: For the diagnosis and classification of patients with axial spondyloarthritis (axSpA) imaging (X-ray and magnetic resonance imaging (MRI)) of the sacroiliac joints (SIJ) is indispensable. The aim of this study was to analyze the impact of clinical information on the radiologists' findings and image interpretation.

Methods: X-ray and MRI images of 109 patients with suspected axSpA were analyzed separately by three radiologists in four consecutive rounds. Initially, only the X-ray images were analyzed, later X-ray and MRI images were analyzed without and with clinical information (demographic data, SpA characteristics, physical activity, pregnancy) about the patients. The radiographs were analyzed according to the modified New York criteria, the MRI images according to inflammatory and/or structural changes of the SIJ. Analyzation was performed separately per radiologist and combined according to the majority principle (≥2/3 radiologists). The compatibility (precision) of the radiological findings with the clinical diagnosis by rheumatologists (reference standard) as well as sensitivity, specificity, PPV, NPV and the agreement of the radiologists were examined. In addition, the radiologists' self-estimated diagnostic confidence and the estimated influence of clinical information on the radiologic evaluation were recorded and evaluated regarding imaging method and clinical information.

Results: Without clinical information, the precision of X-ray findings was 70% compared to 78% with clinical information, and for X-rays and MRI combined 81% and 82%, respectively. For X-ray alone, the sensitivity and specificity of radiologic findings were 51% and 94% without and 60% and 100 % with clinical information while for X-rays and MRI combined, 74% and 90% vs 71% and 98% were calculated, respectively. The diagnostic confidence of radiologists increased from 5.2 ± 1.9 to 6.0 ± 1.7 with clinical information for X-rays, and from 6.7 ± 1.6 to 7.2 ± 1.6 for X-rays and MRI, respectively. Self-estimated diagnostic confidence increased with clinical information for all imaging modalities, while the influence of clinical information on X-ray alone and X-ray and MRI combined was not assessed differently.

Conclusion: With knowledge of clinical information, the precision, specificity, and diagnostic confidence of radiologic evaluation in patient with suggested axSpA improved. These data favor the provision of clinical information to radiologists.

1 Einleitung

In der Diagnostik der axialen Spondyloarthritis (axSpA) spielen radiologische Untersuchungen eine entscheidende Rolle. Bisher ist unklar, wie und in welchem Ausmaß die Auswertung und Interpretation der Bildgebung bei Patient*innen mit dem Verdacht auf axSpA durch klinische Informationen beeinflusst werden. Kürzlich hat eine ASAS-Expertengruppe (s. Abschnitt 1.2.7) internationale Empfehlungen zum Inhalt der klinischen Informationen erarbeitet, die den befundenden Radiolog*innen bei der Frage nach einer möglichen axSpA zur Verfügung gestellt werden sollten (1). Ziel dieser Arbeit ist es, den Einfluss der klinischen Informationen auf die radiologische Befundung und Interpretation zu untersuchen.

1.1 Axiale Spondyloarthritis

1.1.1 Überblick

Die axiale Spondyloarthritis ist eine meist chronisch verlaufende entzündlich-rheumatische Erkrankung, die vorwiegend das Achsenskelett betrifft und zu strukturellen Veränderungen der Wirbelsäule in Form von Knochenneubildungen führen kann. Sowohl die entzündlichen als auch die strukturellen Veränderungen können mit zum Teil erheblichen Schmerzen und Bewegungseinschränkungen sowie Funktionsdefiziten einhergehen.

Ein besonders typisches Symptom ist der entzündliche Rückenschmerz, der sich unter anderem dadurch auszeichnet, dass er sich eher durch Bewegung als durch Ruhe bessert.

Die axSpA zählt zur Gruppe der Spondyloarthritiden (SpA), einer Familie rheumatischer Erkrankungen mit Beteiligung der Sakroiliakalgelenke (SIG), die eine starke Assoziation mit dem MHC-Klasse 1 Molekül HLA-B27 aufweist (2). Etwa 30% der axSpA-Patient*innen haben auch einen Befall der peripheren Gelenke mit einem meist asymmetrischen Verteilungsmuster, wobei vorwiegend die unteren Extremitäten betroffen sind. Darüber hinaus kommt es gehäuft zu Entzündungen des Kreuz-Darmbeingelenks (Sakroiliitis), der Wirbelkörper (Spondylitis), der Sehnenansätze (Enthesitis) und der Finger und Zehen (Daktylitis).

Im Vergleich zur axialen SpA, die vor allem das Achsenskelett betrifft, umfassen die peripheren SpA Krankheitsbilder, die vorwiegend periphere muskuloskelettale Symptome aufweisen. Entitäten aus beiden Gruppen, der axialen und der peripheren SpA, können gemeinsam oder überlappend auftreten. Patienten aus beiden Gruppen können zusätzlich an extra-muskuloskelettalen Manifestationen einer SpA wie an einer Entzündung der

Regenbogenhaut (Uveitis), an Psoriasis vulgaris und/oder chronisch entzündlichen Darmerkrankungen (CED) leiden (3, 4).

Der Begriff axSpA umfasst alle Stadien dieses Krankheitsbildes. Wendet man die ASAS-Klassifikationskriterien auf Basis der modifizierten New York-Kriterien (mNYC) an, lässt sich formal eine röntgenologische axSpA (r-axSpA) von einer nicht-röntgenologischen axSpA (nr-axSpA) differenzieren. Die Grenze zwischen beiden sind durch die mNYC definiert, d.h. Patienten mit definitiven strukturellen Veränderungen in den SIG werden als r-axSpA und die ohne solche Veränderungen als nr-axSpA klassifiziert. Historisch war diese Unterscheidung erforderlich, wenn ein Pharmaunternehmen eine Zulassung nicht nur für die klassische ankylosierende Spondylitis (AS), im deutschsprachigen Raum auch Morbus Bechterew genannt, sondern auch für die nr-axSpA erreichen wollte.
Heute sind fast alle Biologika – also biotechnologisch hergestellte, zielgerichtet wirkende Medikamente – für beide Indikationen zugelassen. Für die Diagnosestellung ist die Unterscheidung zwischen diesen beiden Entitäten wenig sinnvoll – zumal bekannt ist, wie unzuverlässig die Bewertung von definitiven strukturellen Veränderungen in den SIG ist (5). Darüber hinaus ist es rein klassifikatorisch auch möglich, dass bei einem Patienten formal eine nr-axSpA diagnostiziert wird, obwohl an der Wirbelsäule Syndesmophyten nachweisbar sind und dies unter klinischen Aspekten wenig sinnvoll ist. Die beschriebene Unterteilung suggeriert außerdem, dass die nr-axSpA eine frühe Form und die r-axSpA eine späte Form ist - das stimmt aber nur zum Teil -, denn es gibt nicht wenige Patient*innen, überwiegend Frauen, die im Verlauf keine strukturellen Veränderungen entwickeln (6), und umgekehrt auch Männer, die diese bereits in den ersten beiden Jahren nach Erkrankungsbeginn entwickeln (7).

1.1.2 Epidemiologie

Bezüglich der Epidemiologie der SpA wurden unterschiedliche Prävalenzen angegeben, wobei zur Prävalenz der AS die meisten Daten vorliegen. So wird eine die Prävalenz der AS bzw. axSpA mit einer breiten Spanne von 0,1-1,4% angegeben (8, 9).
Bekannt ist, dass Populationen mit hohem Vorkommen des MHC-Klasse-I Antigens HLA-B27 häufiger unter einer SpA leiden (10). Für Europa und die USA wurden HLA-B27 Prävalenzen von 6-9% ermittelt (11, 12). Die Prävalenz der gesamten SpA (axiale und periphere Spondyloarthritis) wurde in Europa mit 0,54% angegeben (9). In den USA wurde eine AS-Prävalenz von 0,52-0,55% und eine SpA-Gesamtprävalenz von 1,0-1,4% erhoben (12), während diese bei den indigenen Völker der nördlichen Arktis mit 1,6% etwas höher lag. HLA-

B27 ist in nördlich gelegenen Ländern häufiger vorzufinden, während das Merkmal in Subsahara-Afrika und Japan sehr selten ist (13).
Das größte Vorkommen von HLA B-27 konnte bei dem nordamerikanischen, indigenen Volk der Haida mit einer Prävalenz in der Bevölkerung von 50% ermittelt werden - hier fand sich eine AS-Prävalenz von 6% (14). Die niedrigsten AS-Prävalenzen wurden in Süd-Ost Asien (0,07%) und in Subsahara-Afrika (0,02%) beschrieben (9).

1.1.3 Pathogenese

Die Pathogenese der SpA ist bisher nicht vollständig geklärt. Bei der axSpA geht man von einem genetischen Anteil im Bereich von 80% aus, davon sind ca. 25% durch die starke Assoziation mit HLA B27-Allelen erklärt (15).
Darüber hinaus sind die Aminopeptidase ERAP-1 und der IL-23-Rezeptor beteiligt (15). Polymorphismen des ersteren sind interessanterweise nur zusammen mit HLA B27 von Bedeutung. Das spricht für eine Beteiligung des Immunsystems bei der Erkrankung, denn dieses Enzym trimmt Peptiden im endoplasmatischen Retikulum, bevor diese über das HLA B27-Molekül CD8+ T-Zellen präsentiert werden (16). Der IL-23-Rezeptor spielt auch bei Psoriasis und CED eine Rolle (17). Durch den Erfolg der Therapie mit Antikörpern gegen TNFα (18) und IL-17 (19) ist klar, dass auch diese Entzündungsbotenstoffe von pathogenetischer Bedeutung sind. Das trifft auch auf die Januskinase (JAK)-Inhibitoren zu (20).
Wichtig für das klinische Erscheinungsbild der Patient*innen mit axSpA ist die Unterscheidung zwischen Entzündung und Knochenneubildung, wobei erstere vorausgeht. Während die strukturellen Veränderungen im SIG nur begrenzt funktionelle Auswirkungen haben, können diese an der Wirbelsäule erheblichen Funktionsdefiziten verursachen. Das nur von einer Minderheit der Betroffenen erreichte Vollbild der Ankylosierung, die sogenannte Bambusstabwirbelsäule, geht mit einem weitgehenden Verlust der Mobilität der Wirbelsäule einher (2).

1.1.4 Klinisches Bild

Der Krankheitsbeginn der axialen Spondyloarthritis liegt für gewöhnlich im dritten Lebensjahrzehnt. Über 90% der Patient*innen erkranken vor dem 45. Lebensjahr (21). Es sind erkranken zwei- bis dreimal mehr Männer als Frauen von der AS betroffen. Bei der nr-axSpA

dagegen ist die Geschlechterverteilung ausgeglichen, tendenziell ist der Frauenanteil sogar höher (22, 23).
HLA-B27 positive Patient*innen erkranken früher an beiden Formen der axSpA (nr-axSpA und AS) als jene ohne HLA-B27 (21, 24). Der Krankheitsverlauf ist bei homozygoten HLA-B27 positiven Patient*innen etwas schwerer als bei heterozygoten (25).

Die Patient*innen mit axSpA leiden häufig an entzündlichem Rückenschmerz (26). Dieser ist charakterisiert durch eine Dauer von mindestens drei Monaten sowie einem Auftreten vor der vierten Lebensdekade. Typischerweise verschlechtern sich die Schmerzen in Ruhe, während körperliche Bewegung und auch eine Therapie mit NSAR häufig zu einer Linderung führen. Außerdem berichten Patient*innen von nächtlichen Schmerzen, besonders in der zweiten Nachthälfte, sowie morgendlicher Steifigkeit von über 30 Minuten. Nicht selten hat der Gesäßschmerz auch alternierenden Charakter, d.h. er wechselt von einer Seite zur anderen. Der entzündliche Rückenschmerz kann auch an anderen Stellen des Achsenskeletts, vor allem der Wirbelsäule in allen drei Segmenten auftreten (27) (s. Tabelle 1).

Charakteristika des entzündlichen Rückenschmerzes
▪ Morgendliche Steifigkeit von > 30 Minuten
▪ Verbesserung der Rückenschmerzen durch Bewegung und Verschlechterung in Ruhe
▪ Nächtliches Erwachen durch Rückenschmerzen in der zweiten Nachthälfte
▪ Langsamer Beginn
▪ Alternierender Gesäßschmerz
▪ Symptombeginn im Alter von < 40 Jahren

Tabelle 1 Charakteristika des entzündlichen Rückenschmerzes.
Modifiziert nach Rudwaleit et al. (28) und Sieper et al. (27).

Pathophysiologische Grundlage der verantwortlichen Entzündungszustände sind häufig Inflammation im Sehnenansatz-Bereich (Enthesitis), die in der Bildgebung als Sakroiliitis, Spondylitis, Spondyloarthritis, Spondylodiszitis und Kostovertebralarthritis und auch an/in Strukturen im Thoraxbereich wie den Kostosternalgelenken auftreten können. Bereits dann, wenn nur die Entzündung überwiegt, aber vor allem auch, wenn strukturelle Veränderungen hinzukommen, können pathologische Befunde bei der körperlichen Untersuchung erhoben

werden, wie z. B. eine verminderte Thoraxexkursionsfähigkeit, eine verringerte Zervikalrotation und eingeschränkte anteriore und laterale Lumbalflexion (2, 22).

In fortgeschrittenen Krankheitsstadien überwiegen oft die strukturellen Veränderungen gegenüber den entzündlichen, die aber lange persistieren bzw. wieder aufflammen können. Je ausgeprägter die strukturellen Schäden, desto eher weist der typische Habitus der Patient*innen auf die Erkrankung hin (u. a. nach vorn gebeugte Haltung durch die thorakale Hyperkyphose). In diesem Zusammenhang ist es wichtig zu betonen, dass bei den allermeisten Patient*innen in den ersten Jahren der Erkrankung kein pathologischer Befund bei der körperlichen Untersuchung erhoben werden kann. Deshalb kommt in der Frühphase der Erkrankung der Magnet-Resonanz-Tomografie (MRT) eine besonders große Bedeutung zu. Mit der MRT-Technik lassen sich sowohl entzündliche als auch strukturelle Veränderungen darstellen (29).

Unter den peripheren Manifestationen der axSpA sind eine Arthritis und/oder Enthesitis mit 30-50% am häufigsten (8, 22).
Die periphere Arthritis tritt als Mono- oder Oligoarthritis auf, betrifft dabei in der Regel die unteren Extremitäten in einem asymmetrischen Verteilungsmuster und führt zu schmerzhaft geschwollenen und geröteten Gelenken.
Die Enthesitis beschreibt eine Entzündung am Übergang von Sehnen oder Bändern zum Knochen, die häufig an der Achillessehne oder Plantarfaszie vorkommt.
Eine weitere periphere Manifestation ist die Daktylitis, die bei 10% der axSpA-Patient*innen auftritt (8). Die Daktylitis beschreibt eine Schwellung eines Fingers oder einer Zehe, welche durch eine Synovitis und/oder Enthesitis bedingt ist.
Die häufigste extra-muskuloskelettale Manifestation ist die Uveitis, die bei bis zu 25% der Patient*innen mit AS vorkommt und typischerweise akut und unilateral auftritt (23).
Weitere extra-muskuloskelettale Manifestationen sind die Psoriasis vulgaris und die chronisch entzündlichen Darmerkrankungen (Morbus Crohn und Colitis ulcerosa), die bei 5-10% der Patient*innen mit axSpA auftreten (23), sowie deutlich seltener eine Aortitis und eine AV-Blockierung.
Patient*innen mit AS weisen ein erhöhtes Risiko für kardiale Komorbiditäten wie Myokardinfarkte und eine arterielle Hypertonie auf. Ferner kann es zu einer pulmonalen und renalen Beteiligung kommen. Außerdem treten gehäuft Erkrankungen des Knochengewebes wie Osteopenie, Osteoporose und Wirbelkörperfrakturen auf (3).

1.2 Diagnostik und Klassifikation der axSpA

1.2.1 Klinische Diagnostik

Die Diagnosestellung der axSpA ist eine klinische Aufgabe. Parameter, die zur Diagnostik herangezogen werden, sind die Anamnese, die klinische Untersuchung, die Laborbefunde, die Bildgebung und die Berücksichtigung möglicher Differenzialdiagnosen. Allerdings gibt es keine diagnostischen Kriterien und keinen allein beweisenden diagnostischen Test. In Nachfolge der mNYC (30) von 1984 wurden 2009 die ASAS-Klassifikationskriterien (31) etabliert (s. Abschnitt 1.2.8).

1.2.2 Anamnese

Bei Patient*innen mit Verdacht auf axSpA sollte im Rahmen der Anamnese erfasst werden, ob die vorhandenen Rückenschmerzen chronisch sind, ob sie vor dem 45. Lebensjahr aufgetreten sind und ob die Charakteristik typisch für den entzündlichen Rückenschmerz ist.

Die Wahrscheinlichkeit der Diagnose axSpA liegt bei Patient*innen mit entzündlichem Rückenschmerz bei bis zu 30%, während die Prävalenz von axSpA unter Patient*innen mit chronischen Rückenschmerzen bei circa 5% liegt (32). Die Frage nach entzündlichem Rückenschmerz bietet daher eine effektive Screening-Maßnahme bei Patienten*innen mit Verdacht auf axSpA (33).

Des Weiteren gibt die Anamnese Auskunft über die Wirksamkeit einer Therapie mit nicht steroidalen Antirheumatika (NSAR). Da axSpA-Patient*innen vor allem in der Frühphase der Erkrankung ein gutes Ansprechen auf die Therapie mit NSAR (34) aufweisen, gehört die Frage nach der Wirksamkeit einer möglicherweise bereits mit diesen Medikamenten durchgeführten Therapie zur Anamnese dazu. Darüber hinaus ist zu erfragen, ob die Patient*innen an einer der o.g. peripheren Arthritiden bzw. Enthesitiden und/oder extra-muskuloskelettalen Manifestationen der axSpA (Uveitis, Psoriasis, CED) erkrankt sind bzw. jemals erkrankt waren. Eine positive Familienanamnese für SpA ist bei Familienmitgliedern ersten und zweiten Grades mit einem erhöhten Erkrankungsrisiko assoziiert (32).

1.2.3 Körperliche Untersuchung und Labordiagnostik

Die körperliche Untersuchung von axSpA-Patient*innen in der frühen Phase der Erkrankung ist in der Regel wenig hilfreich, da zumeist ein normaler Befund erhoben werden kann. In fortgeschrittenen Phasen können allerdings bedingt durch die chronische Entzündung als auch durch die bereits eingetretenen strukturellen Veränderungen pathologische Befunde erhoben werden, da die Wirbelsäulenbeweglichkeit häufig eingeschränkt ist. Hierzu kann man z. B. die anteriore und laterale Mobilität der Wirbelsäule prüfen (Schober-Test), die Zervikalrotation mit einem Goniometer und die Thoraxexkursion mit einem Zentimetermaß. Im Bereich des SIG können Druck- oder Kompressionsschmerzen, das sog. Mennell-Zeichen, erhoben werden.

Zu den für die Diagnostik relevanten Laboruntersuchungen zählt die Bestimmung von HLA-B27. Dies wird heute i.d.R. mit Durchflusszytometrie (FACS) oder Polymerasekettenreaktion (PCR) untersucht, beide Verfahren haben eine hohe Genauigkeit. Die überwiegende Mehrheit der Patient*innen mit AS (80-95%) und axSpA (70-90%) ist HLA-B27 positiv (3, 9). In der deutschen Bevölkerung sind ca. 8% Träger von HLA-B27.

Die Bestimmung der Akute-Phase-Proteine wie das C-reaktive Protein (CRP) und die Blutsenkungsgeschwindigkeit (BSG) geben zum einen Hinweis auf die systemisch-entzündliche Aktivität der Erkrankung, können aber auch einen Hinweis auf das Vorliegen einer axSpA bei der Diagnostik geben. In Querschnittsstudien zeigte sich das CRP nur bei 50%-60% der Patient*innen mit axSpA erhöht, bei nr-axSpA sogar nur bei 30-40%. Die Spezifität lag mit 80% aber deutlich höher (3, 32).

1.2.4 Bildgebung

In der Diagnostik der axSpA nehmen bildgebende Verfahren einen zentralen Platz ein. Grundsätzlich stehen vor allem folgende Verfahren zur Verfügung:

- konventionelles Röntgen der SIG und der Wirbelsäule (WS)
- Computertomografie (CT) der SIG sowie niedrigdosis-CT (engl.: low-dose-CT) der WS
- MRT der SIG und der WS

In der Frühphase der Erkrankung kommt der Bildgebung der SIG die größte und zentrale Bedeutung zu, weil sich hier die entzündliche Aktivität im Achsenskelett mit Abstand am häufigsten manifestiert. Grundsätzlich sind entzündliche und strukturelle Veränderungen in den

SIG zu unterscheiden. Funktionell spielt das SIG vor allem während der Geburt eine wichtige Rolle.

Während sich mit der konventionellen Röntgenuntersuchung und mit der Computertomografie (CT) nur strukturelle Veränderungen nachweisen lassen, geht die MRT darüber deutlich hinaus, da sich mit dieser Technik zusätzlich entzündliche Vorgänge darstellbar sind. Darüber hinaus hat die MRT-Untersuchung den Vorteil, keine ionisierende Strahlung zu verwenden.

Das konventionelle Röntgen sowie die MRT-Untersuchung sind die derzeit empfohlenen Bildgebungsmethoden für die Diagnostik der axSpA, wobei es eine deutliche Tendenz gibt, sich auf die MRT zu beschränken (35, 36).

Die CT-Untersuchung hat zwar eine hohe Genauigkeit in der Detektion von Erosionen, sie findet aber wegen der damit verbundenen Strahlenexposition im klinischen Alltag höchstens eine Anwendung im Rahmen von Begutachtungen. In den letzten Jahren wurden vermehrt Berichte über die gute Performance von „low-dose“ CT in der Darstellung von Wirbelsäulenveränderungen wie Syndesmophyten veröffentlicht (37). In klinischen Studien wurde dieses Verfahren aber bisher noch nicht eingesetzt.

Andere bildgebende Methoden wie die Szintigrafie und die Sonografie werden für die Diagnostik im Achsenskelett nicht empfohlen (38).

1.2.5 Konventionelle Röntgenuntersuchung

Seit Jahrzenten war die konventionelle Röntgenuntersuchung die führende Bildgebung in der Diagnostik der AS und später auch der axSpA (38). Dies beruhte zunächst auf der Tatsache, dass es keine Alternativen gab. Zwischenzeitlich wurde auch die Szintigrafie vermehrt eingesetzt. Seit Mitte der 80er Jahre stand der Nachweis von Strukturveränderungen der SIG im Mittelpunkt der Diagnose und Klassifikation der axSpA.

Die konventionelle Röntgenuntersuchung eignet sich zur Darstellung von knöchernen Veränderungen im SIG, die als Folge der chronischen Entzündung auftreten. Hierbei spielt das Ausmaß der strukturellen Veränderungen eine entscheidende Rolle. Wenn bereits in größerem Ausmaß Erosionen und/oder Ankylosierungen - also die vollständige Verknöcherung eines Gelenkes - vorliegen, ist dies im konventionellen Röntgenbild gut zu erkennen. Sind die knöchernen Veränderungen aber noch nicht ausgeprägt, kommen sie in der nativradiologischen Diagnostik meist nicht ausreichend zur Darstellung. An dieser Stelle verfügen die CT- und MRT-Untersuchungen über eine höhere Sensitivität und Spezifität (35).

Grundsätzlich ist die Zweidimensionalität des konventionellen Röntgenbildes der Schnittbildtechnik (CT und MRT) unterlegen.

Entzündliche Vorgänge können von den genannten Bildgebungen ausschließlich in der MRT dargestellt werden. Mit Röntgen und CT ist das nicht möglich. Daher ist die klassische Röntgenaufnahme zur Detektion von frühen Krankheitsstadien ohne strukturelle Veränderungen nicht geeignet (3).

Im SIG sind röntgenologisch vor allem subchondrale Sklerosierungen, Gelenkspaltverschmälerungen oder -verbreiterungen und knöcherne Erosionen zu erkennen. Diese sind als Folge der chronischen Entzündung mit Knorpel- und Knochendestruktion und sekundären Reparaturmechanismen mit Knochenneubildung zu erklären. Besonders bekannt ist das schon vor 50 Jahren von dem Radiologen W. Dihlmann beschriebene „bunte Bild“ als Ausdruck fortgeschrittener struktureller, teilweise „perlschnurartiger“ SIG-Veränderungen (39). Darüber hinaus dient das konventionelle Röntgenbild auch der Beurteilung von strukturellen Wirbelsäulenveränderungen (40).

Die Beurteilung der Röntgenaufnahmen der SIG im Hinblick auf strukturelle Veränderungen erfolgt anhand der modifizierten New York Kriterien (mNYC) (30). Definitive Veränderungen im SIG liegen demnach vor, wenn Strukturveränderungen ≥ Grad II beidseits vorliegen, oder unilateral eine Strukturveränderung ≥ Grad III nachweisbar ist. Die Graduierung nach den mNYC ist in Tabelle 2 zu sehen.

Definitive strukturelle Veränderungen nach den modifizierten New York Kriterien

Grad 0	Normal, keine Veränderungen
Grad I	Verdächtige, mögliche Veränderungen
Grad II	Minimale SIG-Veränderungen (wenig Sklerose, wenige Erosionen, keine Gelenkspaltveränderung)
Grad III	Moderate bis fortgeschrittene SIG-Veränderungen (ausgeprägte Sklerose, ausgeprägte Erosionen, Gelenkspaltveränderung, partielle Ankylosierung)
Grad IV	Weitgehende Ankylosierung

Tabelle 2 Graduierung der strukturellen SIG-Veränderungen nach den modifizierten New York Kriterien
(modifiziert nach: van der Linden et al. (30)); Abk.: SIG: Sakroiliakalgelenk.

1.2.6 Magnetresonanztomografie (MRT)

Die MRT nimmt aufgrund der Darstellbarkeit von akuten und aktiven entzündlichen Veränderungen einen großen Stellenwert in der Diagnostik der axSpA ein, denn sie ermöglicht

den Nachweis von frühen Entzündungsstadien und dient damit der Erkennung einer axSpA, auch wenn noch keine strukturellen Veränderungen vorliegen.
Hierbei spielen Techniken der Fettsättigung - also die Unterdrückung der aus Fettgeweben gesendeten Signale - sowie die Gabe von Kontrastmittel eine wichtige Rolle. Die am häufigsten verwendete Fettsättigungstechnik ist die STIR- (Short Tau inversion recovery) Sequenz. Diese Sequenzen ermöglichen die Darstellung von sogenannten Knochenmarködemen (KMÖ). KMÖ zeigen sich in T1- (hypointens bzw. dunkel) und T2-gewichteten (hyperintens bzw. hell) Sequenzen unterschiedlich signalintensiv. Unter einer MRT-Sequenz versteht man eine elektromagnetische Pulssequenz, die der Erzeugung von Schnittbildern in einem MRT-Gerät dient.
Neben akut-entzündlichen Prozessen sind mit der MRT auch chronische, strukturelle Veränderungen wie lokale Fettakkumulation, subchondrale Sklerose, Knochenneubildung und/oder Erosionen im SIG darstellbar (3). Diese lassen sich am besten in T1-gewichteten Sequenzen nachweisen (41). Neuere Studien zeigen sogar, dass sich mittels MRT strukturelle Läsionen im SIG mit höherer Sensitivität als mit dem konventionellen Röntgen nachweisen lassen (36, 42, 43).
Die MRT des Achsenskeletts erfolgt in Ganzkörper-Scannern mit einer Feldstärke von 1,0 – 1,5 Tesla. Dabei werden die SIG in einer semi-koronaren Ebene dargestellt, die der Längsachse des Os sacrum entspricht (41).
Eine MRT-Untersuchung ist heute in der Regel indiziert, wenn der Verdacht auf eine axSpA besteht, die klinischen Parameter sowie das konventionelle Röntgen jedoch nicht zur Diagnosesicherung ausreichen (38).
Eine kontrastmittelgestützte Untersuchung ist für die primäre Fragestellung nach axSpA nicht notwendig, da sich die entzündlichen Prozesse in der STIR-Sequenz zuverlässig nachweisen lassen (44).

1.2.7 ASAS-Gruppe

Die Assessment of SpondyloArthritis international Society (ASAS) ist eine internationale Gruppe von Experten auf dem Gebiet der SpA. (www.asas-group.org)
Aufgaben und Ziele der ASAS sind die Unterstützung und Förderung der translationalen und klinischen Forschung sowie die Verbesserung des Wohlbefindens und der Krankheitsverläufe von Patienten*innen mit Spondyloarthritis.
Zu den Mitteln, um dieses Ziel zu erreichen, gehören:

- Steigerung des Bewusstseins für Spondyloarthritis

- Erleichterung der Frühdiagnose
- Entwicklung und Validierung von Bewertungsinstrumenten
- Evaluierung von Behandlungsmodalitäten

1.2.8 Die ASAS-Klassifikationskriterien für axSpA

Im Jahr 2009 hat ASAS vor dem Hintergrund des zunehmenden Einflusses der MRT auf die Diagnosestellung der axSpA die ASAS-Klassifikationskriterien vorgelegt (31). Es sind die ersten Klassifikationskriterien für die axSpA, die neben HLA B27 erstmalig die MRT-Bildgebung berücksichtigen. Dadurch können auch frühe Stadien der Erkrankung erfasst werden, bei denen (noch) keine strukturellen Veränderungen des Achsenskeletts vorliegen.

In den ASAS-Klassifikationskriterien wurden zwei Hauptarme definiert (s. Abbildung 1). Der bildgebende Arm fordert den Nachweis einer Sakroiliitis in der Bildgebung sowie mindestens einen SpA-Parameter. Als radiologischer Nachweis dient dabei entweder eine definitive röntgenologische Strukturveränderung entsprechend der mNYC oder eine aktive Sakroiliitis in der MRT. Letztere wurde definiert als Nachweis von Knochenmarködem (KMÖ), wenn dieses im subchondralen Knochen lokalisiert ist und sehr suggestiv für SpA ist (45).

Der klinische Arm fordert den Nachweis von HLA-B27 sowie mindestens zwei weiteren SpA-Parameter.

Die ASAS-Klassifikationskriterien haben eine Sensitivität von 82,9% bei einer Spezifität von 84,4% (31).

Die Klassifikationskriterien sind nicht mit Diagnosekriterien gleichzusetzen und dienen primär nicht der individuellen Diagnosestellung. Ziel der Klassifikationskriterien ist es, homogene Gruppen zu formieren und klar zu definieren, um diese beispielsweise in Studien einzuschließen (22, 46). Die Anwendung der Klassifikationskriterien sollte erst erfolgen, nachdem die Diagnose der axSpA gestellt wurde (32).

ASAS-Klassifikationskriterien für axSpA

Einschlusskriterien:

- ≥ 3 Monate Rückenschmerzen
- Alter bei Krankheitsbeginn: < 45 Jahre

Bildgebungs-Arm:		Klinischer Arm:
Sakroiliitis in der Bildgebung[1]	ODER	HLA-B27
Plus		Plus
≥ 1 SpA-Parameter		≥ 2 weitere SpA-Parameter

SpA-Parameter:

- Entzündlicher Rückenschmerz
- Arthritis
- Enthesitis
- Uveitis
- Daktylitis
- Psoriasis vulgaris
- CED (M. Crohn, Colitis ulcerosa)
- Gutes Ansprechen auf NSAR
- Positive Familienanamnese für SpA
- HLA-B27
- Erhöhtes CrP

Abbildung 1 ASAS Klassifikationskriterien
Nach Rudwaleit et al. (31); Abk.: ASAS: Assessment of SpondyloArthritis international Society; axSpA: axiale Spondyloarthritis; HLA: humanes Leukozyten-Antigen; SpA: Spondyloarthritis; CED: chronisch entzündliche Darmerkrankung; CrP: C-reaktives Protein; NSAR: nicht-steriodale Antirheumatika. [1]Sakroiliitis in der Bildgebung entsprechend definitiver struktureller Veränderungen nach den mNYC (30) bzw. entsprechend einer Sakroiliitis in der MRT nach den ASAS-Kriterien (45).

1.2.9 Sakroiliitis in der MRT: ASAS-Definitionen

Die ASAS hat im Rahmen der ASAS-Klassifikation für SpA Kriterien für einen positiven MRT-Befund im Sinne einer SpA festgelegt. Das Vorliegen eines KMÖ stellt dabei das wichtigste Zeichen einer aktiven Sakroiliitis dar, wenn es in einer für die SpA typischen Lokalisation im subchondralen Knochen vorliegt, eindeutig erscheint und hochgradig verdächtig für eine SpA ist (45).

Das KMÖ soll an mindestens zwei typischen Lokalisationen auf einem Schnittbild vorliegen oder auf mindestens zwei aufeinanderfolgenden Schichten an gleicher Lokalisation nachweisbar sein (47). Ein KMÖ nach diesen Kriterien konnte auch bei gesunden

Proband*innen beziehungsweise bei Patient*innen mit unspezifischen Rückenschmerzen in 20 - 40% - i. d. R. in geringerem Ausmaß - nachgewiesen werden (42). AxSpa-Patient*innen weisen jedoch insgesamt einen ausgeprägteren Befund an KMÖ auf als Patient*innen mit mechanisch bedingten Rückschmerzen sowie eine Verteilung mit Betonung der hinteren Gelenkabschnitte, während bei mechanisch bedingten Rückenschmerzen vorwiegend die vorderen Gelenkabschnitte betroffen sind (48). Dem entsprechend hat ein vergleichsweise geringes KMÖ keine Vorhersagekraft für künftige strukturelle Veränderungen im SIG, eine große entzündliche Läsion zusammen mit HLA-B27 aber durchaus (49). Darüber hinaus sagt ein positiver, d.h. entzündlicher MRT-Befund das therapeutische Ansprechen auf Biologika wie die TNFα-Antagonisten voraus (50).

Weitere von der ASAS definierte Zeichen einer aktiven Entzündung in den SIG sind eine Kapsulitis, eine Kontrastmittelanreicherung im Gelenkspalt, eine Entzündung unmittelbar an einer Erosion, eine Enthesitis und der Nachweis von intraartikulärer Flüssigkeit. Diese Befunde weisen ebenfalls auf eine axSpA hin. Um die ASAS-Kriterien zu erfüllen, wird aber zusätzlich der Nachweis eines KMÖ gefordert.

Zu den strukturellen Schäden an den SIG, die in der MRT nachweisbar sind, zählen Erosionen, welche in der T1-Wichtung als Defekt im subchondralen Knochen erscheinen und mit einer Verminderung der Kortikalisbreite sowie einer Signalminderung des angrenzenden Knochenmarks einhergehen. Weitere Zeichen struktureller Schäden sind Fettläsionen, welche in der T1-Wichtung als homogene, hyperintense Veränderungen im subchondralen Knochen erkennbar sind und Fettmetaplasien, die in erosionsbedingten Knochenaussparungen lokalisiert sind, sowie Sklerose, Ankylose und knöcherne Aussprossungen (45).

1.2.10 Berlin Algorithmus

Als Hilfestellung zur Diagnostik der axSpA kann der von der ASAS modifizierte Berlin Algorithmus herangezogen werden (51) (s. Abbildung A 3 im Appendix). Dieser sieht als Eingangskriterien, wie auch die ASAS-Klassifikationskriterien, Rückenschmerzen mit einer Dauer von über drei Monaten sowie einen Symptombeginn vor dem 45. Lebensjahr vor. Die Sensitivität und Spezifität dieses Algorithmus liegen mit 78.5% bzw. 79.6% in einem Bereich, in dem ein erheblicher Teil der Patient*innen mit axSpA fälschlicherweise nicht diagnostiziert wird beziehungsweise 20% der gesunden Patient*innen falsch positiv diagnostiziert werden. Die Expert*innenmeinung stellt, wie auch in dieser Arbeit, den Goldstandard in der Diagnosestellung der axSpA dar.

1.2.11 Diagnoseverzögerung

Die Diagnose wird in den meisten Fällen um viele Jahre verzögert gestellt (52). Eine Erhebung von Daten aus Deutschland zeigt eine mittlere Diagnoseverzögerung von 5,7 Jahren (53), frühere Studien sprachen gar von bis zu 9 Jahren Verzögerung (54). Dies wird vor allem mit der Schwierigkeit erklärt, die vergleichbar wenigen Patient*innen mit entzündlichem Rückschmerz wegen axSpA aus der großen Gruppe der Patient*innen mit Kreuzschmerzen herauszufiltern (3).

1.3 Therapie und Verlauf

Die Remission der Erkrankung ist das wichtigste Therapieziel bei der Behandlung der axSpA. Darunter versteht man das Fehlen von klinischen Zeichen der axSpA (extra-muskuloskelettale und muskuloskelettale Manifestationen) sowie negative Entzündungsparameter im Labor (55). Die Remission bedeutet vor allem aber eine deutlich verminderte Krankheitsaktivität. Diese wird mit dem ASDAS (56) oder BASDAI (57) gemessen, wobei bei einem ASDAS < 1,3 von einer inaktiven Erkrankung ausgegangen werden kann. In der Therapie der axSpA wird das „treat-to-target“ Prinzip verfolgt.

Darüber hinaus soll der röntgenologische Progress der Erkrankung, also die zunehmenden strukturellen Veränderungen des Achsenskeletts, die zu einer deutlichen Mobilitätseinschränkung vor allem der WS führen können, verhindert werden.

In der Behandlung der axSpA spielt ein multimodales Therapiekonzept aus pharmakologischen und nicht-pharmakologischen Maßnahmen, wie u.a. Physiotherapie und Patient*innenschulung, eine entscheidende Rolle (3).

Bei der pharmakologischen Therapie kommen in erster Linie NSAR zum Einsatz. Bei hoher Krankheitsaktivität werden Biologika verwendet, darunter TNF- und IL-17-Inhibitoren sowie Hemmer der Januskinasen (JAK) (55).

Nicht mehr verwendet für die Achsenskelettbeteiligung bei SpA werden konventionelle Immunmodulatoren (csDMARDs) wie Methotrexat und Sulfasalazin sowie systemische Steroide.

1.4 Die Rolle der klinischen Information in der radiologischen Beurteilung

Über alle medizinischen Fachgebiete hinweg stellt sich die viel diskutierte Frage, ob klinische Informationen über Patient*innen einen Einfluss auf die Befundung der Bildgebung durch die

Radiolog*innen haben (58, 59). Oder anders formuliert: Ändert sich die Interpretation eines Bildes, wenn der untersuchenden Person die klinischen Fakten vorliegen bzw. nicht vorliegen?

Überweisungen an radiologische Fachärzt*innen sollten üblicherweise Informationen über Patient*innen enthalten, die der befundenden Person ein größeres Verständnis über den klinischen Kontext ermöglichen. Diese Informationen beinhalten unter anderem Daten über Alter und Geschlecht sowie die aktuellen Symptome der Patientin oder des Patienten, über ihre oder seine Krankheitsgeschichte, den klinischen Gesamtzustand, ggf. die Medikation und die Familienanamnese sowie möglicherweise Angaben über die klinisch vermutete Diagnose. Darüber hinaus sollte von der überweisenden Person eine möglichst klare klinische Fragestellung an die Radiolog*innen gerichtet werden, z.B. welcher klinische Aspekt bei der Beurteilung des Bildes besonders berücksichtigt werden sollte. Davon unabhängig müssen die Radiolog*innen allerdings immer eine Gesamtbeurteilung der untersuchten Aufnahmen vornehmen.

Kürzlich hat eine Expert*innengruppe der ASAS Empfehlungen zur inhaltlichen Ausgestaltung von Anforderungen an die Radiologie und der radiologischen Befunde bei Patient*innen mit Verdacht auf axSpA veröffentlicht (1) (s. Abbildung A 4 im Appendix). Neben zehn Empfehlungen für die Befundung wurden sechs Punkte genannt, die Kliniker*innen im Rahmen der Anforderung an die Radiologie erwähnen sollten. Dazu zählen wichtige klinische Informationen wie Alter, Geschlecht und der HLA-B27-Status, Charakter, Dauer und Lokalisation von aktuellen bzw. zurückliegenden Rückenschmerzen. Darüber hin aus sollte erwähnt werden, ob eine physisch belastende Tätigkeit ausgeübt wird und ob bei Frauen eine Geburt zurückliegt. Außerdem sollten, falls vorhanden, Voraufnahmen bereitgestellt und Informationen zur Kontrastmittelverträglichkeit kommuniziert werden. Abschließend sollte die anfordernde Person die Verdachtsdiagnose sowie mögliche Differentialdiagnosen erwähnen und die Absicht der Untersuchung (Diagnosestellung, Messung von Krankheitsaktivität und/oder Therapieansprechen) erklären.

1.5 Fragestellung

Der Einfluss von klinischen Informationen auf die radiologische Befundung von Röntgen- und MRT-Aufnahmen ist bereits Gegenstand früherer und auch der aktuellen Forschung gewesen. (58-60) Die meisten Autoren haben über die Vorteile der Kommunikation von umfangreichen Patient*innen-Informationen an die befundenden Ärzt*innen berichtet (58, 59).

Bezüglich der Diagnostik der axSpA liegen bis zum aktuellen Zeitpunkt keine Daten zu diesem Thema vor. Nichtsdestoweniger haben aber, wie oben erwähnt, Expert*innen der ASAS-Gruppe Empfehlungen für eine optimale Kommunikation zwischen Kliniker*innen und Radiolog*innen ausgesprochen, die eine umfangreiche Bereitstellung von klinischen Informationen im Rahmen der Anforderung von bildgebenden Verfahren vorsehen. Bisher wurde aber nicht untersucht, ob diese Vorgehensweise wirklich von Vorteil für die Qualität der Befunde hat.

Ziel dieser Arbeit ist es, den Einfluss der klinischen Information auf die Befundung von Röntgen- und MRT-Aufnahmen zu untersuchen.

Hierzu sollen die folgenden Fragen beantwortet werden:

1. Welchen Einfluss hat die klinische Information auf die diagnostische Beurteilung der Röntgenaufnahmen durch die Radiolog*innen?
2. Welchen Einfluss hat die klinische Information auf die diagnostische Beurteilung der Kombination von Röntgen- und MRT-Aufnahmen durch die Radiolog*innen?

Darüber hinaus sollen die folgenden Fragen beantwortet werden:

3. Welchen Einfluss hat die klinische Information auf die Röntgen-Befundung hinsichtlich der Bewertung von definitiven Strukturveränderungen entsprechend der mNYC?
4. Welchen Einfluss hat die klinische Information auf die MRT-Befundung hinsichtlich der Beurteilung von aktiv-entzündlichen und strukturellen Veränderungen?
5. Welchen Einfluss hat die klinische Information auf die subjektiv geschätzte diagnostische Sicherheit der Radiolog*innen?
6. Wie groß schätzen die Radiolog*innen den Einfluss der klinischen Information auf ihre diagnostische Beurteilung ein?
7. Wie verhält sich die diagnostische Genauigkeit in Abhängigkeit von der selbst geschätzten diagnostischen Sicherheit der Radiolog*innen?

2 Material und Methoden

In diesem Abschnitt wird beschrieben, welche Methoden angewendet wurden, um die Fragestellung dieser Arbeit zu bearbeiten und welches Material herangezogen wurde. Um den Einfluss der klinischen Information auf die radiologische Bewertung von Röntgen- und MRT-Bildern von Patient*innen unter der Verdachtsdiagnose axSpA zu messen, wurden die gleichen Bilder zunächst ohne und später mit klinischer Information durch drei Radiolog*innen befundet und die Ergebnisse in eine Datenbank eingetragen. Anschließend wurden die Ergebnisse der Befundung statistisch ausgewertet.

2.1 Material

Im Folgenden sollen die in dieser Arbeit verwendeten Materialien beschrieben werden.

2.1.1 Auswahl des Patient*innenkollektivs

Die in dieser Arbeit verwendeten Patient*innendaten wurden im Rahmen der OptiRef-Studie (52) der Rheumatologie am Campus Benjamin Franklin der Medizinischen Klinik für Gastroenterologie, Infektiologie und Rheumatologie der Charité – Universitätsmedizin Berlin (im Folgenden Rheumatologie der Charité genannt) erhoben.

In der OptiRef-Studie wurde vor dem Hintergrund der bekanntermaßen verzögerten Diagnosestellung bei Patientinnen mit axSpA (53) ein neu konzipierter, online verfügbarer Fragebogen (engl. „online self referral tool“) für Patient*innen mit chronischen Rückenschmerzen als Überweisungswerkzeug zu einem rheumatologischen Zentrum erprobt und mit dem sogenannten „Berlin referral tool“ (61) verglichen. Dieser Online-Fragebogen (www.bechterew-check.de) empfiehlt eine Vorstellung in der Rheumatologie der Charité, wenn die Patient*innen angaben, dass ihre Rückenschmerzen mindestens drei Monate angedauert hatten, der Beginn der Rückenschmerzen vor dem 45. Lebensjahr erfolgt war und wenn mindestens ein weiterer Spondyloarthritis-Parameter erfüllt war. Auf dieser Basis (durch den Onlinefragebogen) stellten sich im Zeitraum von September 2016 bis September 2017 insgesamt 180 Patient*innen in der Rheumatologie der Charité vor.

Darüber hinaus standen Daten von 181 Patient*innen zur Verfügung, die mit Verdacht auf axSpA mit Hilfe des „Berlin referral tool“ durch Orthopäd*innen und Allgemeinmediziner*innen an die Rheumatologie der Charité überwiesen worden waren. Das

„Berlin referral tool“ dient ärztlichen Kolleg*innen als Hilfestellung für Patient*innen mit Verdacht auf axSpA und empfiehlt eine Überweisung in die Rheumatologie, wenn folgende Kriterien erfüllt sind:

1. es liegen chronische Rückenschmerzen seit mindestens drei Monaten vor,
2. die Rückenschmerzen haben vor dem 45. Lebensjahr begonnen und
3. es liegt entweder ein entzündlicher Rückenschmerz oder ein positiver HLA-B27-Befund vor oder es wurde eine Sakroiliitis in der Bildgebung festgestellt.

Des Weiteren standen Daten von 92 Patient*innen mit Verdacht auf axSpA zur Verfügung, die ohne eines der beiden obengenannten Überweisungswerkzeuge durch ärztliche Kolleg*innen an die Rheumatologie der Charité überwiesen worden waren.

Zusammen konnten also Daten von 453 Patient*innen ausgewertet werden, die alle von Rheumatolog*innen umfangreich anamnestisch und körperlich untersucht worden waren. Dabei wurden auch demografische Daten zum Alter, zum Zeitpunkt des Beginns der Rückenschmerzen, zur Dauer der Rückenschmerzen, zur Charakteristik eines entzündlichen Rückenschmerzes, zur Anwesenheit und/oder Anamnese von Arthritis, Enthesitis, Daktylitis sowie von extraartikulären SpA-Manifestationen (Psoriasis, IBD, akute anteriore Uveitis), zum therapeutischen Ansprechen auf NSAR und zur Familiengeschichte bezüglich SpA erhoben. Darüber hinaus wurden laborchemisch die Konzentration des C-reaktiven Proteins und der HLA-B27-Status bestimmt. Röntgenaufnahmen der SIG wurden angefertigt und, falls zur Diagnosestellung notwendig, MRT-Untersuchungen durchgeführt. Anschließend wurde die Diagnose axSpA bzw. keine axSpA durch die Rheumatolog*innen gestellt.

Alle in dieser Studie verwendeten Patient*innendaten sind durch das Ethikvotum der OpiRef-Studie abgedeckt. Alle Untersuchungen der OptiRef-Studie erfolgten nach Zustimmung durch die zuständige Ethikkommission der Charité – Universitätsmedizin Berlin (EA4/161/15) und gemäß der Deklaration von Helsinki von 1975.

Insgesamt standen für die vorliegende Arbeit Datensätze von 453 Patient*innen zur Verfügung, die sowohl klinische Informationen als auch Bildmaterialien beinhalteten. Aus diesem Pool wurde in einem nächsten Schritt die Patient*innen ausgewählt, bei denen die folgenden drei Bildmodalitäten der SIG vorlagen:

1. (mindestens) eine konventionelle Röntgenaufnahme (Beckenübersicht in anterior-posterior Projektion oder in Ferguson-Projektion) sowie

2. MRT-Aufnahmen in den Sequenzen T1 und STIR („Short Tau Inversion Recovery“) jeweils in schräg-koronarer Ebene.

Die verwendeten MRT-Sequenzen entsprechen den Empfehlungen der ESSR-Gruppe, die im Jahr 2015 veröffentlich wurden (62).

Nach Durchsicht der Patient*innenbilder mit dem Bildbetrachtungsprogramm „RadiAnt DICOM Viewer“ (www.radiantviewer.com) lag in ca. einem Viertel der Fälle alle drei Bildmodalitäten vor (s. Abbildung 2). In wenigen Fällen existierten auch zusätzliche MRT-Sequenzen und weitere Röntgenaufnahmen. Die Radiolog*innen wurden angehalten diese zu ignorieren und nicht in die Befundung mit einzubeziehen.

Aufgrund fehlender Daten in der Befunderhebung und nicht-evaluierbarer Röntgenaufnahmen konnten im Rahmen der Auswertung der Röntgen-Befundung (Runden 1 und 2) acht Fälle nicht in die Analysen mit einbezogen werden. Bei der kombinierten Röntgen- und MRT-Auswertung (Runden 3 und 4) traf dies auf zehn Fälle zu (s. Abbildung 2).

2.1.2 Radiolog*innen

Die Auswertung der Röntgen- und MRT-Aufnahmen fand durch drei Radiolog*innen (im Folgenden auch Reader genannt) der Charité – Universitätsmedizin Berlin statt, die allesamt eine große Expertise in der Befundung muskuloskelettaler Bildgebung aufweisen: Reader 1 verfügt über sechs, Reader 2 über drei und Reader 3 über zwölf Jahre Erfahrung.

Die Befunderhebung durch die Radiolog*innen erfolgte mittels der über die Webapplikation REDCap (Research Electronic Data Capture) (63, 64) vorgefertigten Befundmasken (s. Abbildung A 1 im Appendix). Zur Sicherstellung der internationalen Transparenz wurden die Befundmasken in englischer Sprache verfasst.

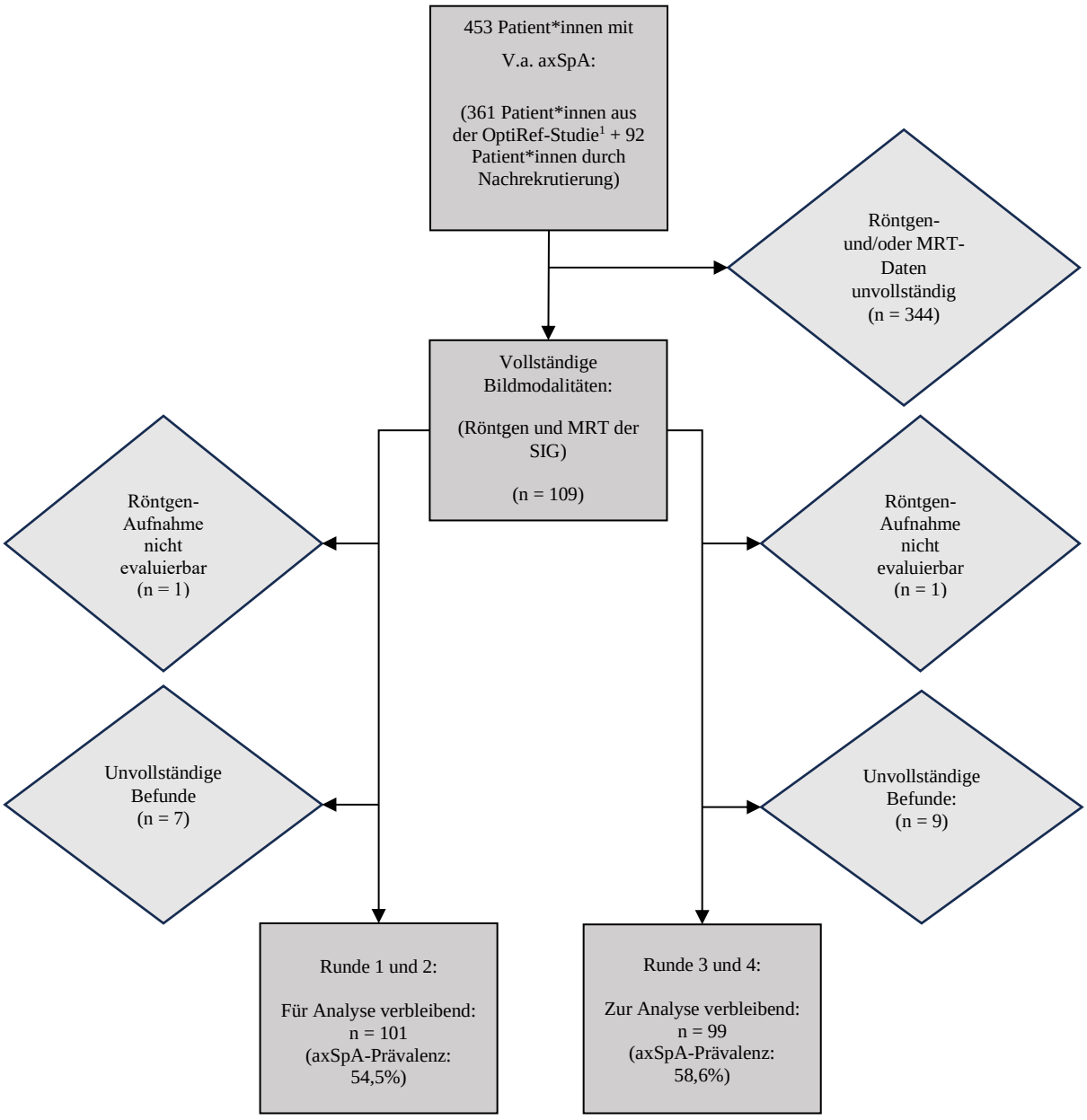


Abbildung 2 Übersicht: Selektion der Studienpopulation
[1]Die Patient*innen-Daten und Bilder entstammen der OptiRef-Studie (52) (Optimal Referral Strategy for Early Diagnosis of Axial Spondyloarthritis); Runde 1: Auswertung der Röntgenbilder ohne klinische Information; Runde 2: Auswertung der Röntgenbilder mit klinischer Information; Runde 3: Auswertung der Röntgen- und MRT-Bilder ohne klinische Information; Runde 4: Auswertung der Röntgen- und MRT-Bilder mit klinischer Information. Die Einschätzung „nicht-evaluierbar“ wurde einstimmig von allen drei beteiligten Radiolog*innen vorgenommen. Unvollständige Befunde: Fälle, die von den Radiolog*innen nicht vollständig befundet wurden. Abk.: MRT: Magnetresonanztomografie; SIG: Sakroiliakalgelenk; axSpA: axiale Spondyloarthritis.

2.2 Radiologische Befundung

Die Auswertung der Röntgen- und MRT-Bilder wurde in vier nacheinander folgenden Runden durchgeführt (s. Abbildung 3). In jeder Runde wurden alle 109 Fälle von jeder Radiologin bzw. jedem Radiologen separat befundet.

Die Reihenfolge der Patient*innen blieb dabei in allen vier Runden unverändert. Nach Abschluss einer Bewertungsrunde wurde die folgende für die befundende Person freigeschaltet. Abgeschlossene Runden waren im Nachhinein für die Reader nicht mehr einsehbar oder veränderbar. Die Befundung der drei Radiolog*innen wurde unabhängig voneinander durchgeführt.

Alle Patient*innen bzw. deren Bilder waren zuvor anonymisiert und mit Identifikationsnummern versehen worden. Die Radiolog*innen hatten keinen weiteren Zugang zur klinischen Information der Patient*innen.

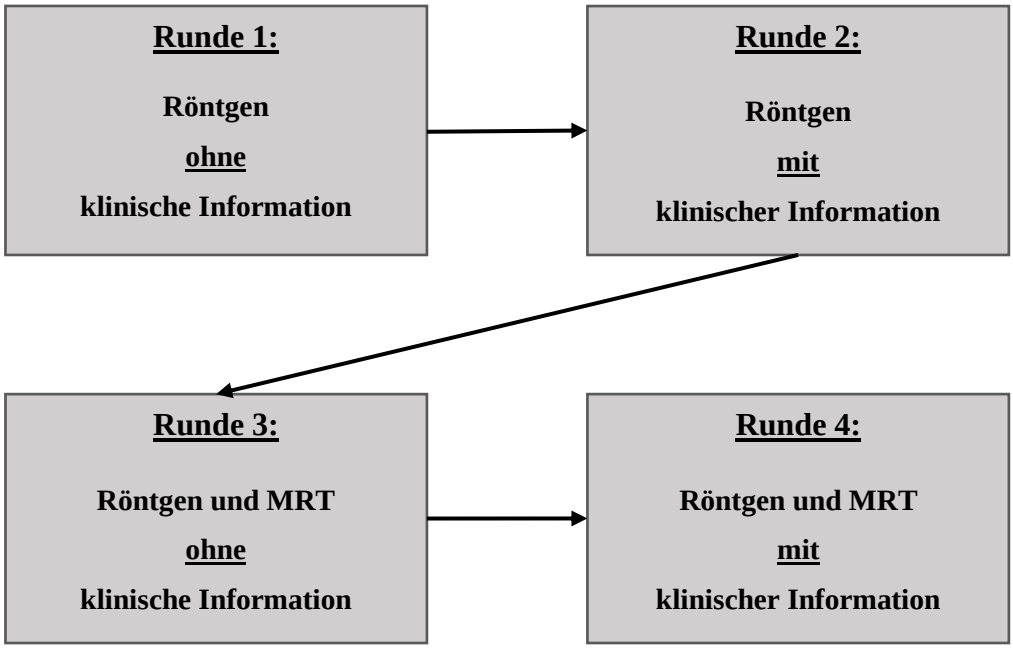


Abbildung 3 Übersicht über die vier Runden der radiologischen Auswertung.
Eine Runde beschreibt die Befundung der 109 Patient*innen-Bilder (Röntgen mit/ohne MRT) durch die Radiolog*innen. Die klinische Information beinhaltet demografische Angaben der Patient*innen sowie u.a. Informationen über Symptome und Laborwerte (s. Methoden). Abk.: MRT: Magnetresonanztomografie.

In der ersten Runde wurden nur die Röntgenaufnahmen der SIG gemäß der Graduierung der mNYC (30) befundet. Das schließt die Graduierung der radiographischen

Gelenkveränderungen in Grad 0 – IV separat für das rechte und linke SIG ein sowie die Möglichkeit, das Bild als nicht-evaluierbar zu bewerten. Außerdem wurde die Frage gestellt, ob der Bildbefund mit der Diagnose einer axSpA übereinstimmt (ja oder nein) und wie hoch die von der beurteilenden Person selbst geschätzte diagnostische Sicherheit bezüglich ihrer Auswertung bewertet wird. Letztere Frage wurde anhand einer numerischen Rating-Skala gemessen, wobei 0 als „nicht sicher“ und 10 als „sehr sicher“ festgelegt wurde. (s. Abbildung A 1 im Appendix). Die klinischen Informationen waren in dieser Runde nicht einsehbar.

In der zweiten Runde der Auswertung wurde den Radiolog*innen zusätzlich zu den Röntgenbildern die jeweils vorhandene klinische Information der Patient*innen vorgelegt (s. Abbildung A 2 im Appendix). Die klinischen Informationen umfassten Alter, Geschlecht, Dauer der Rückenschmerzen, Größe und Gewicht, Lokalisation der Schmerzen (zervikal, thorakal, lumbal, sakral), das Vorliegen eines durch Rheumatolog*innen bestätigten entzündlichen Charakters der Rückenschmerzen sowie Angaben zu anderen SpA-Manifestationen (Uveitis, Psoriasis, CED, Arthritis, Daktylitis, Enthesitis). Darüber hinaus erhielten die Radiolog*innen Informationen zur Familienanamnese bezüglich Spondyloarthritiden, zum Ansprechen auf eine Therapie mit NSAR (definiert als komplette oder deutliche Reduktion der Schmerzen 24 – 48 Stunden nach NSAR-Einnahme in voller Dosis), zur Erwerbstätigkeit (vorwiegend geistige Tätigkeit, vorwiegend körperliche Tätigkeit oder beides), zur sportlichen Aktivität (regelmäßiger Sport ja/nein), zum HLA B27-Status (positiv/negativ) und über die Höhe des CrP-Werts (mg/l).

Außerdem wurde die befundende Person über die Anzahl von Schwangerschaften und Geburten sowie über die Zeitspanne zwischen der letzten Geburt und der Bildaufnahme bei Patientinnen informiert. Die Radiolog*innen erhielten keine Angaben zur gestellten Diagnose. Die Auswertung der Röntgenaufnahmen in Runde 2 erfolgte analog zu den in Runde 1 verwendeten Kriterien, einschließlich der Schätzung der diagnostischen Sicherheit. Zusätzlich wurde auf einer numerischen Rating-Skala angegeben, wie stark die klinische Information die Entscheidung über die radiologische Befundung beeinflusst hat (0: kein Einfluss; 10: sehr starker Einfluss).

In der dritten Runde der Auswertung wurden den Radiolog*innen sowohl die Röntgen- als auch die MRT-Aufnahmen präsentiert, jedoch ohne die zugehörige klinische Information. Die Beurteilung der Röntgen-Aufnahmen fand entsprechend der Kriterien der Runde 1 statt. Die Befundung der MRT-Bilder erfolgte anhand folgender vier Ja-/Nein-Fragen:

1. Gibt es Zeichen von inflammatorischer Aktivität, die mit einer SpA kompatibel sind?
2. Gibt es entzündliche Veränderungen? (einschließlich nicht SpA-typischer Veränderungen)
3. Liegen strukturelle Veränderungen vor, die mit einer SpA kompatibel sind?
4. Sind strukturelle Veränderungen vorhanden? (einschließlich nicht SpA-typischer struktureller Veränderungen)

Auch in Runde 3 wurde die diagnostische Sicherheit der Befundung entsprechend den Runden 1 und 2 subjektiv eingeschätzt.

In der vierten und letzten Runde wurden die Röntgen- und MRT-Aufnahmen bei vorliegender klinischer Information ausgewertet. Wie in Runde 2 sollte auch hier der selbst geschätzte Einfluss der klinischen Information auf der Skala von 0 – 10 angegeben werden.

Die kombinierte Befundung der Röntgen- und MRT-Bilder aus den Runden 3 und 4 sollten dabei die radiologische Befundung im klinischen Alltag simulieren, bei der häufig mehrere Bildmodalitäten simultan berücksichtigt werden.

2.3 Statistische Auswertung

Die statistische Auswertung erfolgte unter Verwendung des Tabellenkalkulationsprogramms Microsoft Excel (Version 2308) sowie des Statistikprogramms SPSS (Version 28.0.1.0).

Für Beschreibung der Charakteristika des Patient*innenkollektivs wurde zunächst eine deskriptive Statistik angefertigt (s. Tabelle 3).

Anschließend wurde die Auswertung der Radiolog*innen beschrieben und die Häufigkeiten der röntgenologischen und in der MRT sichtbaren Veränderungen dargestellt, die wiederum als solche gewertet wurden, wenn eine Übereinstimmung von mindestens zwei der drei Radiolog*innen vorlag (s. Tabelle 4). Dabei wurden die Ergebnisse der Runden verwertet, die für die entsprechende Fragestellung die jeweils höchste Interrater-Reliabilität aufwiesen, d.h. für die Röntgen-Befunde die Runde 2 und für die Röntgen/MRT-Auswertung die Runde 4. Aufgrund der o.g. fehlenden Daten liegen bei sechs axSpA Patient*innen aus der Röntgen/MRT-Runde keine Ergebnisse bezüglich definitiver Strukturveränderungen nach den mNYC aus Runde 2 vor. Von den 55 axSpA-Patient*innen aus der Röntgen-Bewertung fehlen drei MRT-Befunde aus der Runde 4.

Die Hauptfragestellung, welchen Einfluss das Vorliegen von klinischen Informationen auf die diagnostische Genauigkeit durch die Radiolog*innen hat, wurde separat zum einen für die rein röntgenologische Diagnostik (Runden 1 und 2) und zum anderen für die bildgebende Diagnostik aus der Kombination von Röntgen und MRT (Runden 3 und 4) untersucht. Dabei wurden die Übereinstimmungen der von den Radiolog*innen als ‚mit der Diagnose axSpA kompatibel' bzw. ‚nicht mit der Diagnose axSpA kompatibel' beurteilten Fälle mit dem vordefinierten Goldstandard ermittelt und prozentual dargestellt, dies wird im Folgenden „Präzision" genannt. Als Goldstandard diente die klinische Diagnose, die durch die behandelnden Rheumatolog*innen gestellt wurde.

Die Entscheidung der Radiolog*innen für die Kompatibilität mit der Diagnose axSpA wurde als solche gewertet, wenn mindestens zwei der drei Radiolog*innen den Befund als kompatibel beurteilt hatten. Wenn umgekehrt weniger als zwei von drei Beurteiler*innen das vorliegende Bild als kompatibel mit der Diagnose axSpA beurteilt hatten, wurde dies als „nicht kompatibel mit der Diagnose axSpA" gewertet.

Anschließend erfolgte der Vergleich der Präzision aus der Runde ohne klinische Information mit der Runde mit klinischer Information.

Bei der Fallzahlberechnung wurde angenommen, dass ohne klinische Informationen eine Diskrepanz von 20 % zwischen der diagnostischen Beurteilung der Radiolog*innen und der endgültigen Diagnose durch die Rheumatolog*innen zu beobachten ist. Mit klinischer Diagnose wurde eine Abweichung der diagnostischen Urteile von 5% erwartet. Es werden mindestens 85 Bilder benötigt, um den statistisch signifikanten Unterschied zwischen den Ergebnissen ohne und mit klinischen Informationen mit Hilfe des McNemar-Tests mit einer Power von 80 % und einem Signifikanzniveau von $\alpha = 0{,}05$ zu ermitteln. Zur Analyse der Inter-Rater-Reliabilität wurde Fleiss' Kappa berechnet.

Darüber hinaus wurden Sensitivität, Spezifität, der positive prädiktive Wert (PPV) sowie der negative prädiktive Wert (NPV) berechnet.

Diese Analysen wurden ebenso für jeden Reader einzeln vorgenommen und dargestellt.

Zu der Fragestellung, welchen Einfluss das Vorliegen klinischer Informationen auf die Beurteilung der Röntgenaufnahmen anhand der modifizierten New York Kriterien hatte, wurden die Ergebnisse der Runde eins (ohne klinische Informationen) mit denen der Runde zwei (mit klinischen Informationen) verglichen, d.h. in wie vielen Fällen war das röntgenologische Kriterium der mNYC von mindestens zwei der drei Radiolog*innen als erfüllt

angesehen worden. Das röntgenologische Kriterium der mNYC gilt als positiv, wenn bilateral eine Sakroiliitis von Grad II - IV oder unilateral eine Sakroiliitis von Grad III - IV vorliegt. Als statistischer Test wurde der McNemar-Test eingesetzt, zur Inter-Rater-Reliabilitätsanalyse wurde Fleiss‘ Kappa berechnet. Darüber hinaus wurden diese Analysen auch für jeden Reader separat durchgeführt.

Zu der Fragestellung, welchen Einfluss das Vorliegen der klinischen Information auf die Beurteilung der MRT-Aufnahmen bezüglich der Feststellung von aktiv entzündlichen und strukturellen Veränderungen (jeweils SpA-typische bzw. unspezifische Veränderungen) hat, wurden die Ergebnisse aus den Runden drei und vier verglichen. Wenn mindestens zwei der drei Reader in einer MRT-Aufnahme die gleiche Veränderung festgestellt hatten, wurde diese Veränderung zusammenfassend positiv, d.h. als vorliegend bewertet. Es konnten pro Fall mehrere Veränderungen gleichzeitig als vorliegend bewertet werden (Mehrfachbenennungen möglich, z.B. SpA-typische und nicht SpA-typische strukturelle Läsionen vorhanden). Als statistischer Test wurde der McNemar-Test eingesetzt, zur Inter-Rater-Reliabilitätsanalyse wurde Fleiss‘ Kappa berechnet. Diese Analysen wurden ebenso für jeden Reader separat durchgeführt.

Die Ergebnisse der Frage nach der von den Radiolog*innen selbst geschätzten diagnostischen Sicherheit sowie die Ergebnisse der Frage nach dem subjektiv geschätzten Einfluss der klinischen Information wurden separat für jeden Reader deskriptiv dargestellt (s. Tabellen 10 – 13).

Darüber hinaus wurde eine deskriptive Darstellung als Abbildung (Boxplot) aus den jeweiligen Mittelwerten der drei Reader sowohl für die selbst geschätzte diagnostische Sicherheit als auch für den selbst geschätzten Einfluss durch die klinische Information angefertigt (s. Abbildungen 7-9). Zur Veranschaulichung der Auswertung der einzelnen Radiolog*innen wurden Balkendiagramme angefertigt, die die jeweilige Änderung der selbst geschätzten diagnostischen Sicherheit (Differenz aus den Werten vor und nach Kenntnisnahme der klinischen Information) vergleichend mit dem subjektiv geschätzten Einfluss der klinischen Information eines jeden Readers darstellen (s. Abbildung 10 und Abbildung 11).

Des Weiteren wurde dargestellt, in wie vielen Fällen es bei vorliegender klinischer Information im Vergleich zur Befundung ohne klinische Information zu einer Abnahme bzw. zu einer Zunahme der selbst geschätzten diagnostischen Sicherheit kam. Veränderungen der selbst geschätzten diagnostischen Sicherheit wurden unabhängig vom Ausmaß der Abweichung

gewertet. Es wurde geprüft, in wie vielen Fällen mindestens zwei der drei Reader eine Erhöhung bzw. Senkung der subjektiv geschätzten diagnostischen Sicherheit angegeben hatten. Fälle, bei denen dies nicht zutraf, wurden als nicht verändert bewertet. Diese Analysen wurden sowohl für die rein röntgenologische Diagnostik (Runden 1 und 2) als auch für die gesamte Bildgebung, d.h. mit MRT (Runden 3 und 4) durchgeführt sowie für jeden Reader separat (s. Tabellen 14 und 15).

Zuletzt wurde die Übereinstimmung des radiologischen Befundes mit dem diagnostischen Goldstandard (Präzision) in Abhängigkeit von der subjektiv geschätzten diagnostischen Sicherheit in drei Kategorien (niedrig [0 – 3], moderat [4 – 6], hoch [7 – 10]) untersucht und prozentual dargestellt. Diese Kategorisierung erfolgte anhand der in den Runden ohne klinische Information (Runden 1 bzw. 3) geschätzten diagnostischen Sicherheit und wurde für die entsprechende Runde mit klinischer Information (Runden 2 bzw. 4) beibehalten. Berechnet wurden jeweils die Präzision, die durchschnittliche diagnostische Sicherheit, die Sensitivität, die Spezifität, der PPV und der NPV sowie in den Runden 2 bzw. 4 der Mittelwert des selbst geschätzten Einflusses durch die klinische Information. Diese Analysen wurden jeweils mit und ohne klinische Information und sowohl für jeden einzelnen Reader als auch auf der Gruppenebene durchgeführt. Bei letzterer wurde die diagnostische Genauigkeit ebenfalls nach dem Mehrheitsprinzip (mindestens zwei von drei Readern) bewertet und die Kategorisierung nach der selbst geschätzten diagnostischen Sicherheit anhand der jeweiligen Mittelwerte der drei Reader vorgenommen. Dabei wurde ab einem Wert von 3,5 bzw. 6,5 aufgerundet (s. Tabellen 16 – 19 sowie Abbildung 12 und Abbildung 13).

3 Ergebnisse

3.1 Demografische Daten und klinische Befunde

Die demografischen Daten aller eingeschlossen 109 Patient*innen sind in Tabelle 3 dargestellt. Insgesamt hatten 61 Patient*innen (56%) die Diagnose axSpA erhalten, bei 48 Patient*innen hatte keine axSpA vorgelegen. Im Vergleich hatte es in der Gruppe mit axSpA mehr Männer (64%) gegeben, die Patient*innen waren jünger (35 vs. 38 Jahre) und die Symptomdauer war etwas kürzer gewesen (7 vs. 8 Jahre). Die Angaben hinsichtlich des Vorliegens eines entzündlichen Rückenschmerzes wiesen eine deutliche Diskrepanz auf. So lag der entzündliche Rückenschmerz in der in der Gruppe mit axSpA bei 90% gegenüber 50% in der non-axSpA-Gruppe vor. Während Uveitis und Psoriasis in der axSpA-Gruppe häufiger vorkamen, war dies bei CED anders. Eine CED war in beiden Gruppen nur bei wenigen Patient*innen zu erheben gewesen. Während niemand über eine Daktylitis berichtete hatte, war dies bei peripherer Arthritis und Enthesitis anders: hier hatte es in der non-axSpA-Gruppe mehr Patient*innen mit Anhalt für die beiden SpA-Manifestationen gegeben.

In der axSpA-Gruppe hatten mehr Patient*innen eine positive Familienanamnese für SpA angegeben (25% vs. 4%) und auch HLA B27 war in einem deutlich höheren Prozentsatz nachgewiesen worden (85% vs. 26%). Ein erhöhter CrP-Wert war bei axSpA-Patient*innen geringfügig häufiger zu erheben gewesen (28% vs. 27%). Deutlich häufiger war in der axSpA-Gruppe über ein gutes Ansprechen auf eine Therapie mit NSAR berichtet worden (80% vs. 57%). Höchstens geringfügige Unterschiede zeigten sich im Prozentsatz der Patient*innen, die vermehrte körperliche Tätigkeiten angegeben hatten (26% vs. 25%). AxSpA-Patient*innen waren weniger sportlich aktiv gewesen (56% vs. 65%). In der axSpA-Gruppe hatten 41% der Frauen angegeben, jemals schwanger gewesen zu sein, während dies in der non-axSpA-Gruppe auf 64% der Frauen zutraf.

Demografische und klinische Charakteristika der Patient*innen

	Gesamt (n = 109)	axSpA[1] (n=61)	Non-axSpA[1] (n = 48)
Diagnose axiale SpA[1], n (%)	61 (56.0%)		
männliches Geschlecht, n (%)	54 (49.5%)	39 (63.9%)	15 (31.3%)
Alter in Jahren, *M* (*SD*)	36.4 (10.5)	34.9 (10.9)	38.3 (9.8)
Dauer Rückenschmerz, in Jahren, *M* (*SD*)	7.4 (6.5)	6.7 (6.5)	8.2 (6.6)
Entzündlicher Rückenschmerz[2], n (%)	79 (72.5%)	55 (90.2%)	24 (50%)
Uveitis / Iritis, jemals: n (%)	12 (11.0%)	11 (18%)	1 (2.1%)
Psoriasis, jemals: n (%)	8 (7.3%)	6 (9.8%)	2 (4.2%)
CED, jemals, n (%)	3 (2.8%)	1 (1.6%)	2 (4.2%)
Periphere Arthritis, jemals: n (%)	18 (16.5%)	9 (14,8%)	9 (18.8%)
Daktylitis, jemals, n (%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
Enthesitis, jemals, n (%)	19 (17.4%)	8 (13.1%)	11 (22.9%)
Positive Familienanamnese SpA, n/N (%)	15/98 (15.3%)	13/53 (24.5%)	2/45 (4.4%)
Positiver HLA-B27 Status, n/N (%)	61/104 (58.7%)	49/58 (84.5%)	12/46 (26.1%)
CrP-Wert erhöht (> 5mg/l), n (%)	30 (27.5%)	17 (27.9%)	13 (27.1%)
Gutes Ansprechen auf NSAR[3], n/N (%)	61/87 (70.1%)	40/50 (80.0%)	21/37 (56.8%)
Vorwiegende körperliche Tätigkeit, n (%)	28 (25.7%)	16 (26.2%)	12 (25.0%)
Regelmäßiger Sport, n (%)	65 (59.6%)	34 (55.7%)	31 (64.6%)
Schwangerschaft, jemals, n/N (weibl. Geschlecht) (%)	30/55 (54.5%)	9/22 (40.9%)	21/33 (63.6%)

Tabelle 3 Demografische und klinische Charakteristika der Patient*innen
[1]Diagnose gestellt durch Rheumatolog*innen; [2]Entzündlicher Rückenschmerz bewertet durch Rheumatolog*innen Abk.: axSpA: axiale Spondyloarthritis; non-axSpA: keine axSpA-Diagnose; M: Mittelwert; SD: Standardabweichung; CED: chronisch entzündliche Darmerkrankung; SpA: Spondyloarthritis; HLA: Humanes Leukozytenantigen; NSAR: nicht-steroidale Antirheumatika;

3.2 Ergebnisse der radiologischen Befundung

Bei der Auswertung der radiologischen Befundung zeigt sich, dass - in Übereinstimmung von mindestens zwei von drei Readern - definitive röntgenologische Strukturveränderungen (nach den mNYC) bei 33 der 55 Patient*innen mit axSpA vorliegen. Dies entspricht einem Prozentsatz von 60% in der axSpA-Gruppe. Zwischen den Readern wurden dabei Werte zwischen 56.4 % (Reader 1) – 67.3 % (Reader 3) angegeben.

Diese Daten wurden aus den Ergebnissen der Runde 2 (mit klinischer Information) erhoben, in der die Interrater-Übereinstimmung der Röntgenbefunde besser war als in der Runde 1.

Die kombinierte Röntgen- und MRT-Auswertung der Runde 4 ergibt, dass sich die Radiolog*innen mit einer Mehrheit von mindestens zwei von drei Readern darüber einig waren, dass bei 66% der axSpA-Patient*innen SpA-typische aktiv-entzündliche Veränderungen nachzuweisen waren und bei 67% SpA-typische strukturelle Veränderungen. In 62% der Fälle

mit axSpA lagen beide Veränderungen gleichzeitig vor. Dagegen wurde in der non-axSpA-Gruppe angegeben, dass SpA-typische aktive Entzündung und Strukturveränderungen bei jeweils 2% der Patient*innen zu sehen waren. Bei 29% der axSpA-Patient*innen wurde angegeben, dass weder aktiv-entzündliche noch strukturelle Veränderungen, die SpA-typischen Charakter haben, nachgewiesen werden konnten. Diese Daten wurden aus der Runde 4 erhoben, in der die Interrater-Reliabilität bezüglich der Frage nach SpA-typischen Veränderungen im MRT am größten war. (s. Tabelle 4)

	Alle Patient*innen[1]:		**Nur axSpA-Patient*innen[2,3]: (n= 55)**	
	axSpA[3] (n=58)	non-axSpA[3] (n=41)	mNYC positiv[4] (n=33)	mNYC negativ[5] (n=22)
Röntgen[7]: mNYC positiv[4], % (n/N)	59.6% (32/52)	0% (0/41)	100%	0%
MRT[8]:				
- Aktive Entzündung[6], % (n/N)	65.5%	2.4%	87.1% (27/31)	38.1% (8/21)
- Strukturveränderung[6], % (n/N)	67.2%	2.4%	96.8% (30/31)	33.3% (7/21)
- Entzündung und Strukturveränderung[6], % (n/N)	62.1%	2.4%	87.1% (27/31)	33.3% (7/21)
- Keine SpA-typischen Veränderungen, % (n/N)	29.3%	97.6%	3.2% (1/31)	61.9% (13/21)

Tabelle 4 Häufigkeiten SpA-typischer radiologischer Veränderungen
Röntgenologische Strukturveränderungen und Spondyloarthritis(SpA)-typische Veränderungen im MRT. Wenige Daten fehlen hier, dies ist in den Methoden beschrieben. [1]basierend auf den Daten der Runde 4 (n=99); [2]Patient*innen mit der klinischen Diagnose axiale Spondyloarthritis (axSpA) aus Runde 2 (n=55); [3]klinische Diagnose durch Rheumatolog*in; [4]definitive röntgenologische Strukturveränderungen in den Sakroiliakalgelenken (SIG) nach den modifizierten New York Kriterien (mNYC) (30); [5]keine definitiven röntgenologischen Strukturveränderungen in den SIG nach den mNYC; [6]SpA-typische Veränderungen. [7]basierend auf der radiologischen Auswertung der Runde 2; [8]basierend auf der radiologischen Auswertung der Runde 4; Weitere Abk.: MRT: Magnetresonanztomografie.

Nimmt man als Standard alle Patient*innen, die entweder im Röntgen und/oder in der MRT einen positiven Befund für SpA-typische strukturelle Veränderungen hatten, finden sich unter den 52 axSpA-Patient*innen 31 mNYC positive (59.6%) und 37 MRT positive (71.2%). Die Diskrepanz zwischen den in erster Zeile angegebenen 55 axSpA-positiven Patient*innen und den 52 zur Berechnung herangezogenen axSpA-Patient*innen kommt durch die drei fehlenden MRT-Befunde aus den Runden der Röntgen-Befundung zustande.

Von allen axSpA-Patient*innen, die entweder im Röntgen oder MRT einen Nachweis von Strukturveränderungen im SIG zeigten (n = 38), wurden mittels Röntgen 31 als solche erkannt (81.6%) während dies bei der MRT 37 (97.4%) waren. Dies entspricht einer Verbesserung von

18.4% in der Erkennung von Strukturveränderungen mittels MRT-Befundung im Vergleich zur konventionellen Röntgenuntersuchung.

3.3 Auswertung der radiologischen Befundung: Die diagnostische Genauigkeit

Der erste primäre Endpunkt-Parameter war die Präzision der Radiolog*innen in der Befundung der Röntgenaufnahmen der SIG im Vergleich mit und ohne vorliegender klinischer Information. Die diesbezüglichen Ergebnisse sind in Tabelle 5 dargestellt. Hierbei zeigt sich, dass die Präzision des Befundes deutlich besser war, wenn klinische Informationen vorlagen (70.3 vs. 78.2 % ($p < 0.008$)). Auch alle anderen Parameter wurden durch das Vorliegen der klinischen Information eher verbessert, wobei die Sensitivität und der negativ-prädiktive Wert (NPV) eher moderat waren (60% und 68%), während die Spezifität und der positiv-prädiktive Wert (PPV) 100% erreichten. Der PPV lag damit deutlich über dem NPV (100% vs. 68%) Die Interrater-Reliabilität (Kappa-Werte) steigen von einer moderaten Übereinstimmung aus der Runde ohne klinische Information auf eine substanzielle Übereinstimmung an (0.56 vs. 0.66) (65).

	Röntgen ohne klinische Information (Runde 1)	**Röntgen mit klinischer Information** (Runde 2)	**p-Wert[2]:**
Präzision[1]:	70.3% (71/101)	78.2% (79/101)	0 .008
Sensitivität:	50.9%	60%	
Spezifität:	93.5%	100%	
PPV:	90.3%	100%	
NPV:	61.4%	67.7%	
Interrater-Reliabilität[2]:	$\kappa = 0.56$	$\kappa = 0.66$	
Prävalenz (axSpA):	54.5%		

Tabelle 5 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen-Befundung

Aggregierte Daten der drei Radiolog*innen. [1]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); [2]Interrater-Reliabilität berechnet mittels Fleiss' Kappa. [2]p-Wert berechnet mittels McNemar-Test. Abk.: PPV: positiv-prädiktiver Wert, NPV: negativ-prädiktiver Wert; axSpA: axiale Spondyloarthritis.

Der zweite primäre Endpunkt war die Präzision der radiologischen Befunde unter Berücksichtigung aller vorliegenden bildgebenden Modalitäten (d.h. Röntgen- und MRT-Aufnahmen). Hierbei fand sich bei Vorliegen der klinischen Information nur eine geringe Zunahme der Präzision (81% vs. 82%), während die Sensitivität eher abnahm (74% vs. 71%), die Spezifität jedoch von 90% auf nahezu 98% anstieg. Der PPV verbesserte sich moderat (92% vs. 98%) und der NPV nahm geringfügig ab (71% vs. 70%) (s. Tabelle 6).

Wie bei der alleinigen Röntgenauswertung liegen die PPV deutlich über den NPV. Die Interrater-Reliabilität erreichte mit und ohne klinische Information eine substanzielle Übereinstimmung (0.77 vs. 0.75) und liegt deutlich höher als bei der alleinigen Röntgen-Auswertung.

	Röntgen und MRT ohne klinische Information (Runde 3)	**Röntgen und MRT mit klinischer Information** (Runde 4)	**p-Wert[2]:**
Präzision[1]:	80.8% (80/99)	81.8% (81/99)	1.00
Sensitivität:	74.1%	70.7%	
Spezifität:	90.2%	97.6%	
PPV:	91.5%	97.6%	
NPV:	71.2%	70.2%	
Interrater-Reliabilität[2]:	$\kappa = 0.77$	$\kappa = 0.75$	
Prävalenz (axSpA):	58.6%		

Tabelle 6 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen- und MRT-Befundung.
Aggregierte Daten der drei Radiolog*innen. [1]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); [2]Interrater-Reliabilität berechnet mittels Fleiss' Kappa. [2]p-Wert berechnet mittels McNemar-Test. Abk.: PPV: positiv-prädiktiver Wert, NPV: negativ-prädiktiver Wert; axSpA: axiale Spondyloarthritis.

In Abbildung 4 sind die Veränderungen nach Kenntnisnahme der klinischen Informationen vergleichend zwischen der alleinigen Röntgenauswertung und der Auswertung der gesamten verfügbaren Bildinformationen dargestellt. Man sieht zum einen, dass Veränderungen im Bereich unter 10% stattgefunden haben. Bei der alleinigen Röntgenbildbewertung zeigt sich bei

allen Werten eine deutlichere Veränderung und Verbesserung für Sensitivität, Spezifität, und der prädiktiven Werte als bei der Auswertung von Röntgen- und MRT-Bildern. Die Änderungen bei Röntgen und MRT sind dagegen zum Teil gegensinnig, d.h. die Sensitivität nimmt ab, während die Spezifität zunimmt.

Während sich die Inter-Rater-Reliabilität beim Röntgen allein durch die klinische Information deutlich verbessert, zeigt sie bei der Auswertung von Röntgen und MRT kombiniert keine wesentliche Veränderung (Δκ +0.08 vs. -0.02).

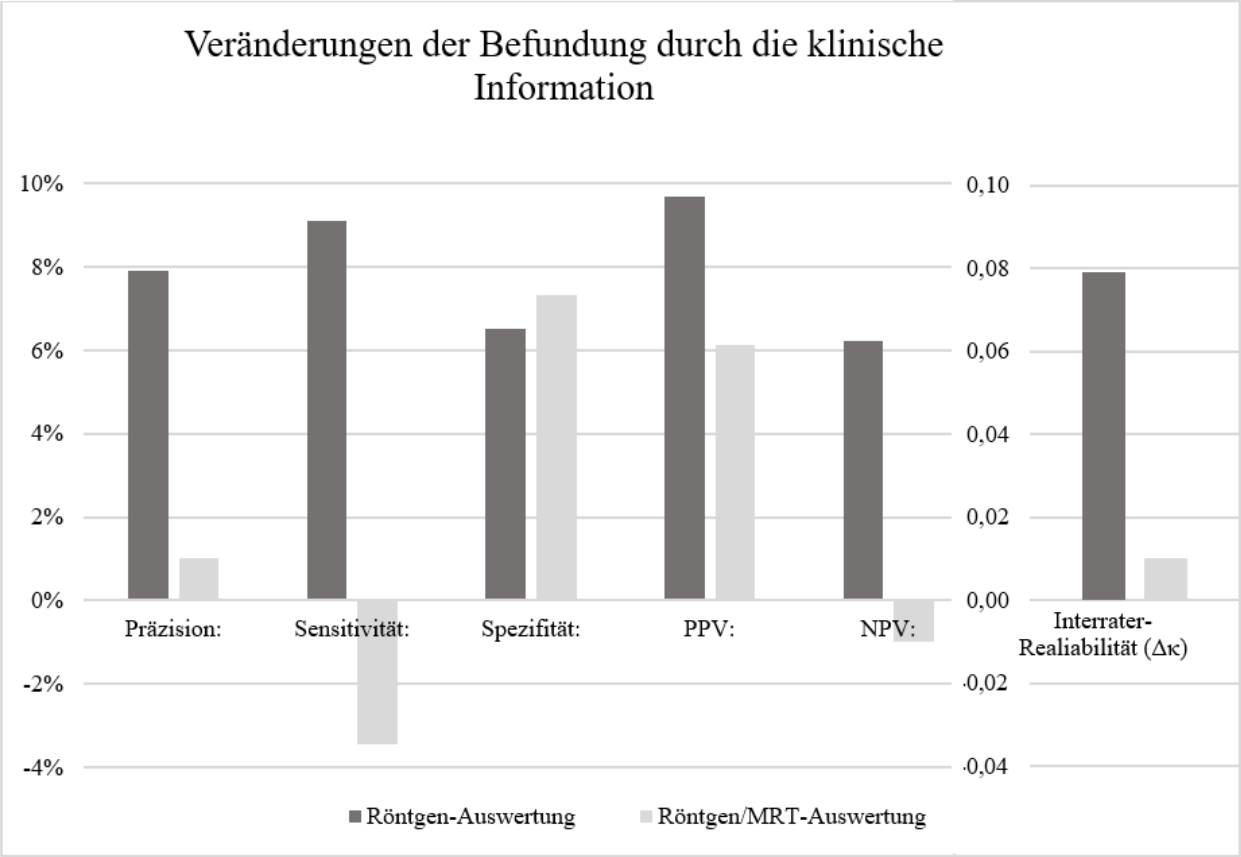


Abbildung 4 Veränderungen der Befundung durch die klinische Information
Aggregierte Daten der drei Radiolog*innen. Präzision: Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen; Abk.: PPV: positiv-prädiktiver Wert, NPV: negativ-prädiktiver Wert; axSpA: axiale Spondyloarthritis. Interrater-Reliabilität berechnet mittels Fleiss' Kappa.

Tabelle 7 zeigt die Variabilität der Röntgen-Befundung der drei Radiolog*innen mit einer Präzision von 63% - 76% ohne klinische Information gegenüber 64% - 77% mit klinischer Information in der Röntgenauswertung. Die Unterschiede der Röntgenbefunde mit und ohne klinische Information waren bei Reader 1 und 2 eher gering ausgeprägt. Eine deutliche Veränderung zeigen dagegen die Ergebnisse von Reader 3, der ohne klinische Information eine Sensitivität der Röntgen-Befunde von nur 38% aufwies, während diese mit gegebener klinischer Information auf 62% anstieg. Im Gegensatz dazu nahm die Sensitivität bei den anderen beiden Readern etwas ab (Reader 1: 64% vs. 56% und Reader 2: 55% vs. 51%).

	Röntgen ohne klinische Information (Runde 1)			**Röntgen mit klinischer Information** (Runde 2)		
	Reader 1	**Reader 2**	**Reader 3**	**Reader 1**	**Reader 2**	**Reader 3**
Präzision[1]:	76.2%	63.4%	63.4%	76.2%	64.4%	77.2%
Sensitivität:	63.6%	54.6%	38.2%	56.4%	50.9%	61.8%
Spezifität:	91.3%	73.9%	93.5%	100%	80.5%	95.7%
PPV:	89.7%	71.4%	87.5%	100%	75.7%	94.4%
NPV:	67.7%	57.6%	55.8%	65.7%	57.8%	67.7%

Tabelle 7 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen-Befundung (separat pro Reader)
[1]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); Abk.: PPV: positiv-prädiktiver Wert, NPV: negativ-prädiktiver Wert; axSpA: axiale Spondyloarthritis.

Bei der Auswertung der Röntgenbilder zusammen mit den MRT-Aufnahmen (Runden 3 und 4) zeigt sich eine etwas geringere Variabilität der Präzision der einzelnen Reader von 77% - 83% ohne klinische Information gegenüber 78% - 84% bei vorliegender klinischer Information, während bei allen Readern ein geringfügiger Zuwachs der Präzision zu verzeichnen ist (s. Tabelle 8).

	Röntgen und MRT ohne klinische Information (Runde 3)			**Röntgen und MRT mit klinischer Information** (Runde 4)		
	Reader 1	**Reader 2**	**Reader 3**	**Reader 1**	**Reader 2**	**Reader 3**
Präzision[1]:	79.8%	76.8%	82.8%	81.8%	77.8%	83.8%
Sensitivität:	72.4%	79.3%	72.4%	72.4%	74.1%	72.4%
Spezifität:	90.2%	73.2%	97.6%	95.1%	82.9%	100%
PPV:	91.3%	80.7%	97.7%	95.5%	86%	100%
NPV:	69.8%	71.4%	71.4%	70.9%	69.4%	71.9%

Tabelle 8 Diagnostische Genauigkeit in Abhängigkeit der klinischen Information der Röntgen- und MRT-Befundung (separat pro Reader)

[1]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); Abk.: PPV: positiv-prädiktiver Wert, NPV: negativ-prädiktiver Wert; axSpA: axiale Spondyloarthritis.

3.4 Befundung der Röntgenbilder nach den mNYC

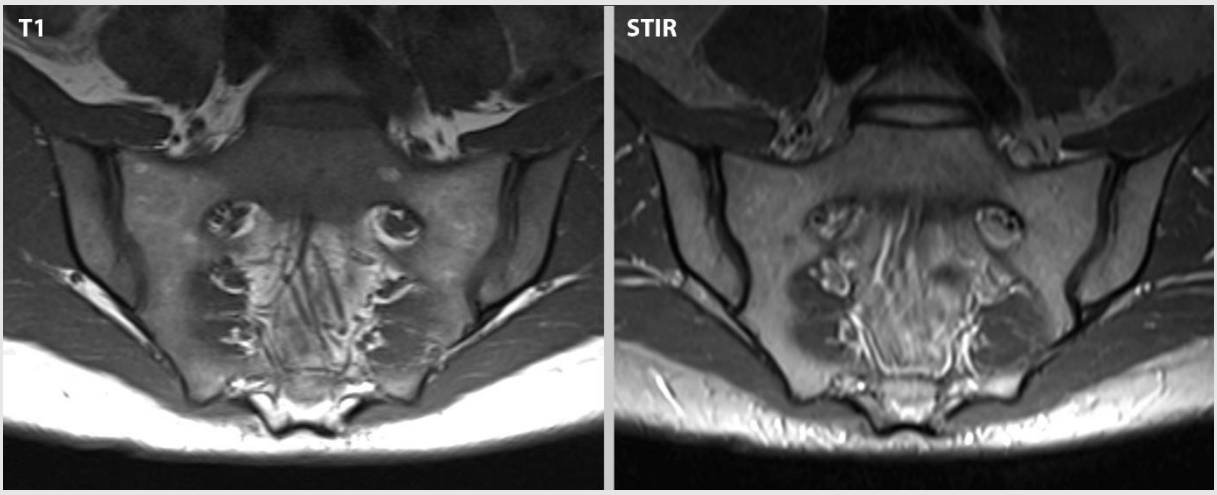
Abbildung 5 Strukturelle Röntgenveränderungen der SIG
Beckenübersichtsaufnahmen (Ferguson-Projektion). Grad 0: Normalbefund der Sakroiliakalgelenke (SIG), der Gelenkspalt ist beidseits gut einsehbar und glatt begrenzt. Grad 2: mittel- bis langstreckige Sklerosierungen beidseits; Grad 3: fortgeschrittene Sklerosierung beidseits und Erosionen mit „Pseudodilatation des Gelenkspalts“. Grad 4: weitgehende Ankylosierung der SIG.

Die Frage nach der Einschätzung der Radiolog*innen bezüglich der Erfüllung der mNYC im Vergleich der Röntgen-Auswertung mit und ohne klinische Information zeigte sich, dass mit klinischer Information bei 33% aller Fälle definitive Strukturveränderungen entsprechend der mNYC beobachtet wurden, während dies ohne klinische Information nur bei 28% der Fall war

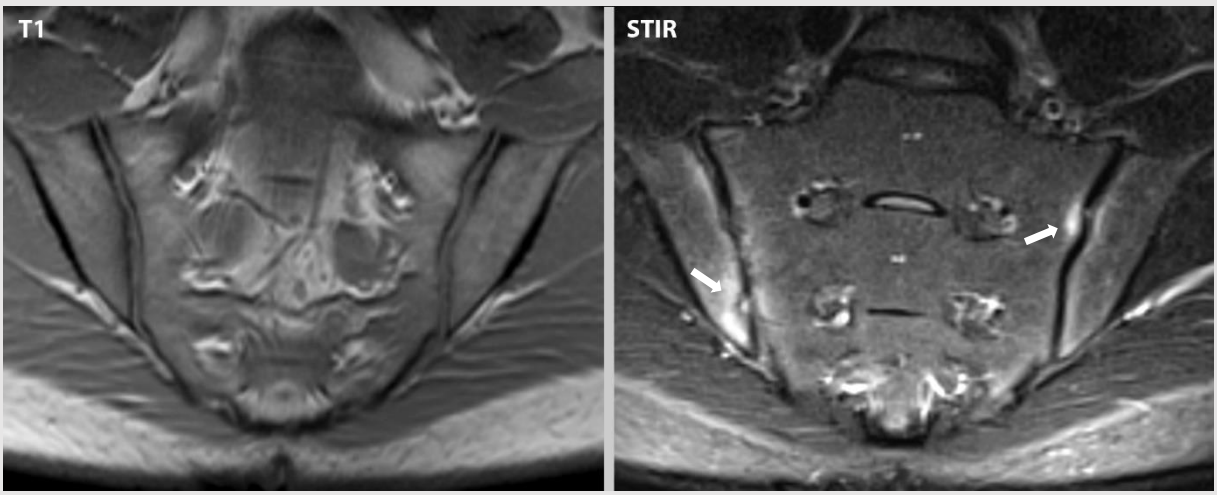
(p = 0.18). Die Interrater-Reliabilität war in beiden Runden substanziell (0.61 ohne und 0.66 mit klinischer Information).

Bei den Ergebnissen der einzelnen Reader zu dieser Frage zeigte sich eine deutliche Variabilität der Fälle, die ohne klinische Information als positiv im Sinne der mNYC gewertet wurden (24% - 36%,), während die Variabilität zwischen den Readern bei vorliegender klinischer Information auf 31% - 37% eher abnahm.

3.5 Befundung der MRT-Aufnahmen: SpA-typische Veränderungen

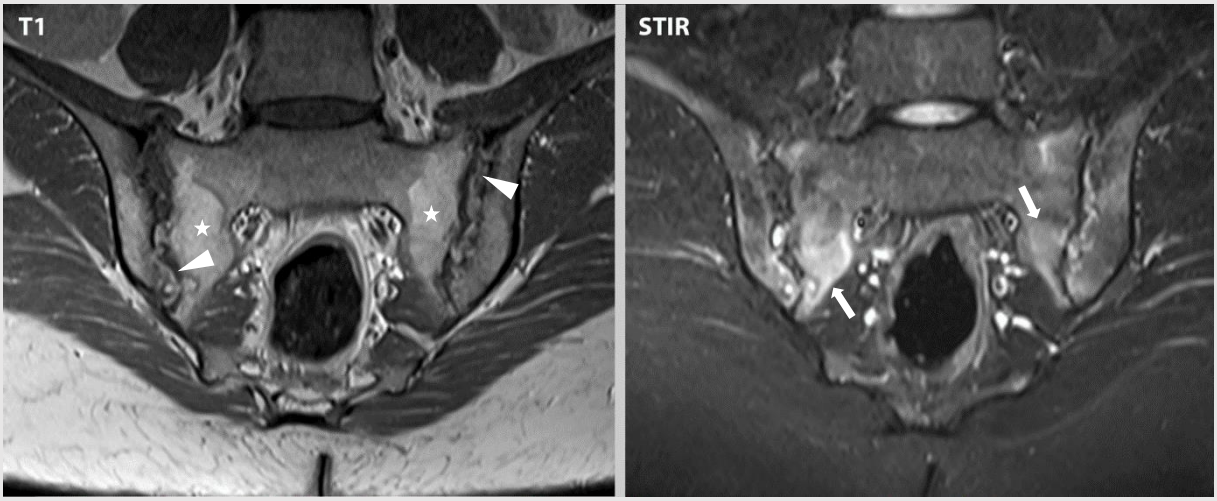


Normalbefund der Sakroiliakalgelenke (SIG) im MRT.



Aktiv entzündliche Veränderungen der SIG ohne strukturelle Läsionen. Die T1-Sequenz zeigt keine strukturellen Auffälligkeiten, während in der STIR-Sequenz hyperintense Signale entsprechend eines subchondralen Knochenmarködems imponieren (Pfeile).

Aktiv entzündliche Veränderungen der SIG, in der STIR-Sequenz sichtbar als hyperintenses Signal (Pfeil). In der T1-Sequzen zeigen sich erosive Veränderungen des Gelenkspalts (Pfeilspitzen).



Fortgeschrittene Sakroiliitis. In der T1-Sequenz zeigen sich ausgeprägte Gelenkspaltsveränderungen (Pfeilspitzen) sowie subchondrale Fettläsionen (Sternchen). In der STIR-Sequenz zeigt sich ein deutliches Knochenmarködem (Pfeile).

Abbildung 6 axSpA typische SIG-Veränderungen in der MRT
Die Aufnahmen erfolgten in schräg-koronarer Ebene. Aufnahmesequenzen: T1-Sequenz und STIR (short tau inversion recovery).

Beim detaillierten Vergleich der MRT-Befunde hinsichtlich SpA-typischen entzündlichen und strukturellen Veränderungen in den SIG waren die Kappa-Werte in einem sehr guten Bereich (> 0.74) und dies sowohl in der Auswertung mit und ohne klinische Information.

Bei den Häufigkeiten dieser Auswertung zeigte sich, dass mit klinischer Information weniger SpA-typische aktive Entzündungen (46% vs. 39%; p = 0.07) als auch signifikant weniger SpA-typische strukturelle Veränderungen (48% vs. 40% p = 0.039) befundet wurden (s. Tabelle 9).

Auch in der separaten Auswertung der drei Reader zeigt sich dieser Trend.

	Röntgen und MRT ohne klinische Information (Runde 3)	**Röntgen und MRT mit klinischer Information** (Runde 4)	**p-Wert[2]**
Aktive Entzündung[1]	45.5% (κ = 0.743)	39.4% (κ = 0.751)	0.07
Strukturelle Läsionen[1]	47.5% (κ = 0.784)	40.4% (κ =0.780)	0.039
Prävalenz axSpA	58,58%		

Tabelle 9 Aktive Entzündung und strukturelle Läsionen (MRT-Befundung) Aggregierte Daten aus der MRT-Befundung der drei Radiolog*innen. [1]aktiv entzündliche bzw. strukturelle Veränderungen, die als axSpA-spezifisch bewertet wurden; [2]p-Werte berechnet durch McNemar-Test; Abk. axSpA: axiale Spondyloarthritis

Die Ergebnisse der Frage nach unspezifischen Veränderungen in den MRT-Aufnahmen – aktiv-entzündliche oder strukturell – zeigten im Vergleich mit und ohne klinische Information jeweils keinen relevanten Unterschied.

3.6 Die subjektiv geschätzte diagnostische Sicherheit

In der Auswertung der Röntgenaufnahmen (Runden 1 und 2) war mit Kenntnis der klinischen Information eine Zunahme der subjektiv geschätzten diagnostischen Sicherheit zu verzeichnen (Mediane: 5.0 vs. 6.0) (s. Abbildung 7).

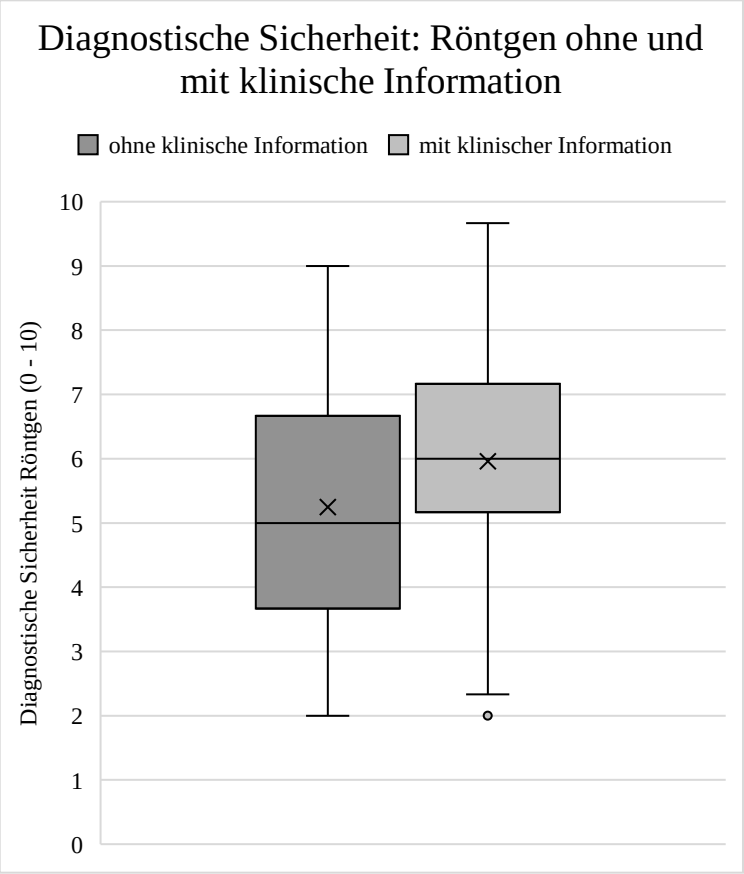


Abbildung 7 Diagnostische Sicherheit: Röntgen
Selbst geschätzte diagnostische Sicherheit (0 = sehr unsicher; 10 = sehr sicher) in der Röntgen-Auswertung im Vergleich zwischen „ohne klinische Information“ und „mit klinischer Information“; Aggregierte Daten der drei Radiolog*innen. x: Mittelwerte.

In der detaillierten Auswertung der einzelnen Reader zeigte sich, dass bereits ohne klinische Information unterschiedliche Angaben zur selbst geschätzten diagnostischen Sicherheit gemacht wurden (Bereich 4.6 – 5.9). Darüber hinaus war der Zuwachs der diagnostischen Sicherheit unterschiedlich stark ausgeprägt (Bereich +0.2 – +1.3). Während der Mittelwert von Reader 2 von einer diagnostischen Sicherheit von 4.6 ohne klinische Information auf 4.8 zunahm, steigt dieser bei Reader 1 von 5.3 auf 5.9 und bei Reader 3 sogar von 5.9 auf 7.2. Verwendet man die Mediane dieser Ergebnisse, so sind die Unterschiede weniger stark ausgeprägt (s Tabellen 10 und 11).

		Reader 1	Reader 2	Reader 3
Mittelwert		5.3	4.6	5.9
Std.-Abweichung		1.9	1.9	3.3
Median		6	5	7
Minimum		1	1	1
Maximum		9	9	10
Perzentile	25	4	3	2
	50	6	5	7
	75	7	6	9

Tabelle 10 Diagnostische Sicherheit: Röntgen ohne klinische Information (separat pro Reader)
Selbst geschätzte diagnostische Sicherheit (0 = sehr unsicher; 10 = sehr sicher)

		Reader 1	Reader 2	Reader 3
Mittelwert		5.9	4.8	7.2
Std.-Abweichung		2.2	2.1	2.4
Median		7	5	7
Minimum		1	1	1
Maximum		10	9	10
Perzentile	25	4	3	6
	50	7	5	7
	75	7.5	6	9

Tabelle 11 Diagnostische Sicherheit: Röntgen mit klinischer Information (separat pro Reader)
Selbst geschätzte diagnostische Sicherheit (0 = sehr unsicher; 10 = sehr sicher)

Die mittlere subjektiv geschätzte diagnostische Sicherheit liegt bei der Auswertung der Runden mit Röntgen- und MRT-Bildern bereits ohne klinische Information über der mittleren diagnostischen Sicherheit der alleinigen Röntgen-Auswertung (Mittelwerte: 6.7 vs. 5.2). In der Runde 4, also nach Kenntnisnahme der klinischen Information, steigt die subjektiv geschätzte diagnostische Sicherheit an (Mediane 6.7 vs. 7.7) (s. Abbildung 8).

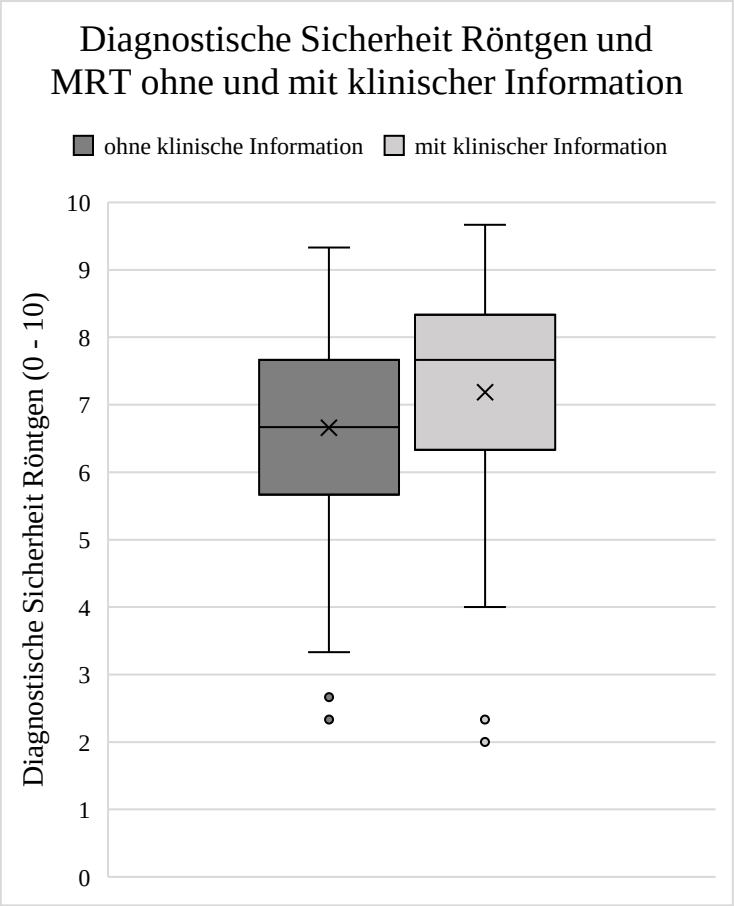


Abbildung 8 Diagnostische Sicherheit: Röntgen- und MRT-Auswertung
Selbst geschätzte diagnostische Sicherheit (0 = sehr unsicher; 10 = sehr sicher) in der Röntgen- und MRT-Auswertung im Vergleich zwischen „ohne klinische Information“ und „mit klinischer Information“; Aggregierte Daten der drei Radiolog*innen. x: Mittelwerte.

In der differenzierten Betrachtung der drei Radiolog*innen aus den Runden 3 und 4 (Röntgen- und MRT-Bildgebung) zeigte sich wie bei der alleinigen Auswertung der Röntgenaufnahmen, dass die drei Reader ihre diagnostische Sicherheit bereits ohne klinische Information im Mittel unterschiedlich einschätzten (Bereich 6.0 – 7.4). Bei Reader 1 und Reader 3 stieg die selbst geschätzte diagnostische Sicherheit, während sie bei Reader 2 leicht sank (s. Tabellen 11 und 12).

		Reader 1	Reader 2	Reader 3
Mittelwert		6.54	6.00	7.44
Std.-Abweichung		1.49	2.13	2.87
Median		8	6	10
Minimum		1	1	1
Maximum		9	10	10
Perzentile	25	7	4	9
	50	8	6	10
	75	8	7	10

Tabelle 12 Diagnostische Sicherheit Röntgen und MRT ohne klinische Information (separat pro Reader)
Selbst geschätzte diagnostische Sicherheit (0 = sehr unsicher; 10 = sehr sicher).

		Reader 1	Reader 2	Reader 3
Mittelwert		6.92	5.89	8.75
Std.-Abweichung		2.01	2.09	2.44
Median		7	6	9
Minimum		2	1	1
Maximum		8	10	10
Perzentile	25	6	4	6
	50	7	6	9
	75	8	7	10

Tabelle 13 Diagnostische Sicherheit Röntgen und MRT mit klinischer Information (separat pro Reader)
Selbst geschätzte diagnostische Sicherheit (0 = sehr unsicher; 10 = sehr sicher).

Bei den Ergebnissen der Befundung der Röntgenbilder (Runde 1 und 2) zeigt sich, dass es in 48% der Fälle zu einer Zunahme der selbst geschätzten diagnostischen Sicherheit kam, wenn die klinische Information vorlag. Dagegen nahm diese bei knapp einem Viertel der Fälle (24%) ab. In 29% blieb die selbst geschätzte diagnostische Sicherheit unverändert (s. Tabelle 14).

Diagnostische Sicherheit[1]:	Röntgen	Röntgen + MRT
Sinkend, n (%):	24 (23.8%)	7 (7.1%)
Unverändert, n (%):	29 (28.7%)	41 (41.4%)
Steigend, n (%):	48 (47.5%)	51 (51.5%)

Tabelle 14 Änderungen der diagnostischen Sicherheit durch klinische Information (nur Röntgen und Röntgen + MRT) - aggregierte Daten der drei Reader
Änderungen der selbst geschätzten diagnostischen Sicherheit bei vorliegender klinischer Information im Vergleich zur Befundung ohne klinische Information. Röntgen (Runde 1 und 2) und Röntgen- und MRT-Auswertung (Runden 3 und 4). [1]selbst geschätzt. n: Anzahl der Fälle.

In der detaillierten Auswertung der einzelnen Reader bezüglich der Veränderungen der selbst geschätzten diagnostischen Sicherheit zeigten sich keine größeren Unterschiede (s. Tabelle 15).

Dagegen zeigt sich bei der Röntgen- und MRT-Auswertungen aus den Runden 3 und 4, dass die selbst geschätzte diagnostische Sicherheit in nur 7% der Fälle sank, während sie bei 52% nach Kenntnis der klinischen Information höher eingeschätzt wurde. In 41% der Fälle blieb diese unverändert (s. Tabelle 14).

Auch hier ergaben sich in der Einzelauswertung der drei Reader keine wesentlichen Unterschiede. (s. Tabelle 15).

	Diagnostische Sicherheit[1]	Reader 1	Reader 2	Reader 3
Röntgen	Sinkend, n (%):	31 (30.7)	34 (33.7)	31 (30.7)
	Unverändert, n (%):	13 (12.9)	27 (26.7)	20 (19.8)
	Steigend, n (%):	57 (56.4)	40 (39.6)	50 (49.5)
Röntgen	Sinkend, n (%):	19 (19.2)	33 (33.3)	9 (9.1)
und MRT	Unverändert, n (%):	26 (26.3)	32 (32.3)	34 (34.3)
	Steigend, n (%):	54 (54.6)	34 (34.3)	56 (56.6)

Tabelle 15 Änderungen der diagnostischen Sicherheit durch klinische Information (separat pro Reader)
Änderungen der selbst geschätzten diagnostischen Sicherheit bei vorliegender klinischer Information im Vergleich zur Befundung ohne klinische Information. Röntgenauswertung (Runden 1 und 2) und Röntgen- und MRT-Auswertung (Runden 3 und 4). Separate Darstellung pro Radiolog*in (Reader). [1]selbst geschätzt. n: Anzahl der Fälle.

3.7 Der subjektiv geschätzte Einfluss der klinischen Information auf die Befundung

Der selbst geschätzte Einfluss der klinischen Information auf die Beurteilung im Vergleich zwischen der Auswertung der Röntgenaufnahmen und der Auswertung der Röntgen- und MRT-Aufnahmen ist in der Abbildung 9 dargestellt. Bei der alleinigen Auswertung der Röntgenaufnahmen zeigt sich eine breitere Streuung der Werte als dies bei der Auswertung der Röntgen- und MRT-Aufnahmen der Fall ist. Die Mittelwerte beider Gruppen liegen auf einem vergleichbaren Niveau.

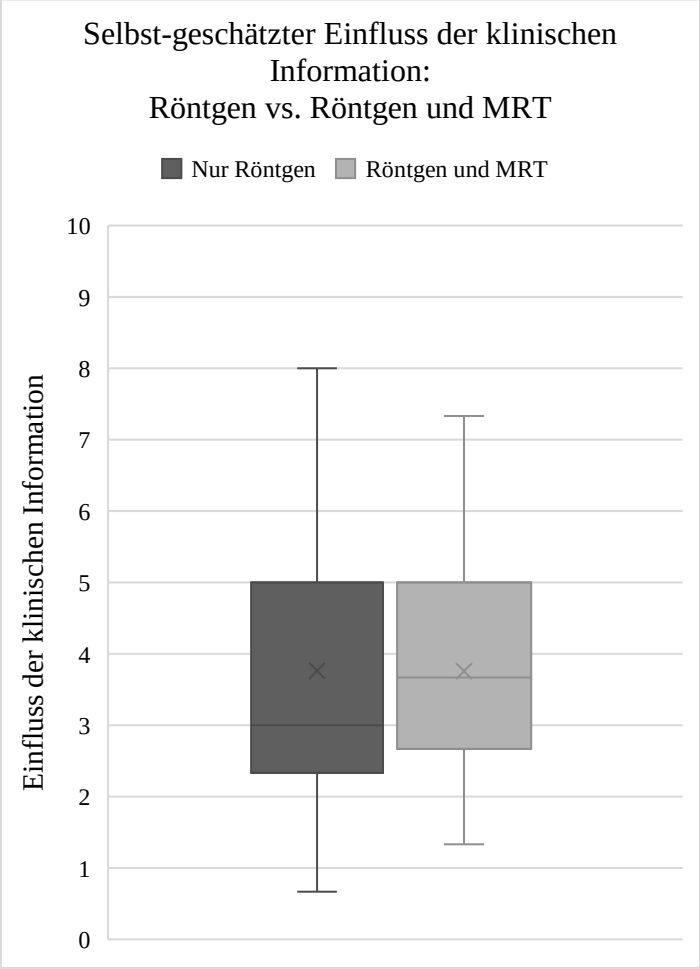


Abbildung 9 Einfluss der klinischen Information Röntgen vs. Röntgen und MRT
Selbst geschätzter Einfluss der klinischen Information der Radiolog*innen auf die Befundung der Röntgenbilder im Vergleich zur Befundung der Röntgen- und MRT-Aufnahmen. 0 = gar kein Einfluss; 10 = maximaler Einfluss. x = Mittelwerte.

In Abbildung 10 und Abbildung 11 sind – jeweils pro Reader – die Veränderungen der selbst geschätzten diagnostischen Sicherheiten neben den Veränderungen der diagnostischen

Genauigkeit (Präzision) sowie dem jeweils eingeschätzten Einfluss der klinischen Information dargestellt. Dabei zeigt sich, dass der Einfluss der klinischen Information von den Readern unterschiedlich eingeschätzt wird (Röntgen: 3.2 – 4.3; Röntgen und MRT: 2.3 – 5).

Während Reader 1 und 2 bei der Röntgen-Befundung einen größeren Einfluss der klinischen Information sehen, zeigt sich bei ihnen nur eine geringe Auswirkung auf die diagnostische Sicherheit und die diagnostische Genauigkeit (Δ Präzision: 0 – 1%). Bei Reader 3 dagegen wird der Einfluss geringer eingeschätzt. Bei ihm zeigt sich aber eine stärkere Verbesserung der Präzision (Δ Präzision: +14%) und auch der selbst geschätzten diagnostischen Sicherheit (+1.3).

Bei der Röntgen und MRT-Auswertung wird der Einfluss der klinischen Information ebenfalls unterschiedlich eingeschätzt, wobei Reader 3 wieder den geringsten Wert angibt.

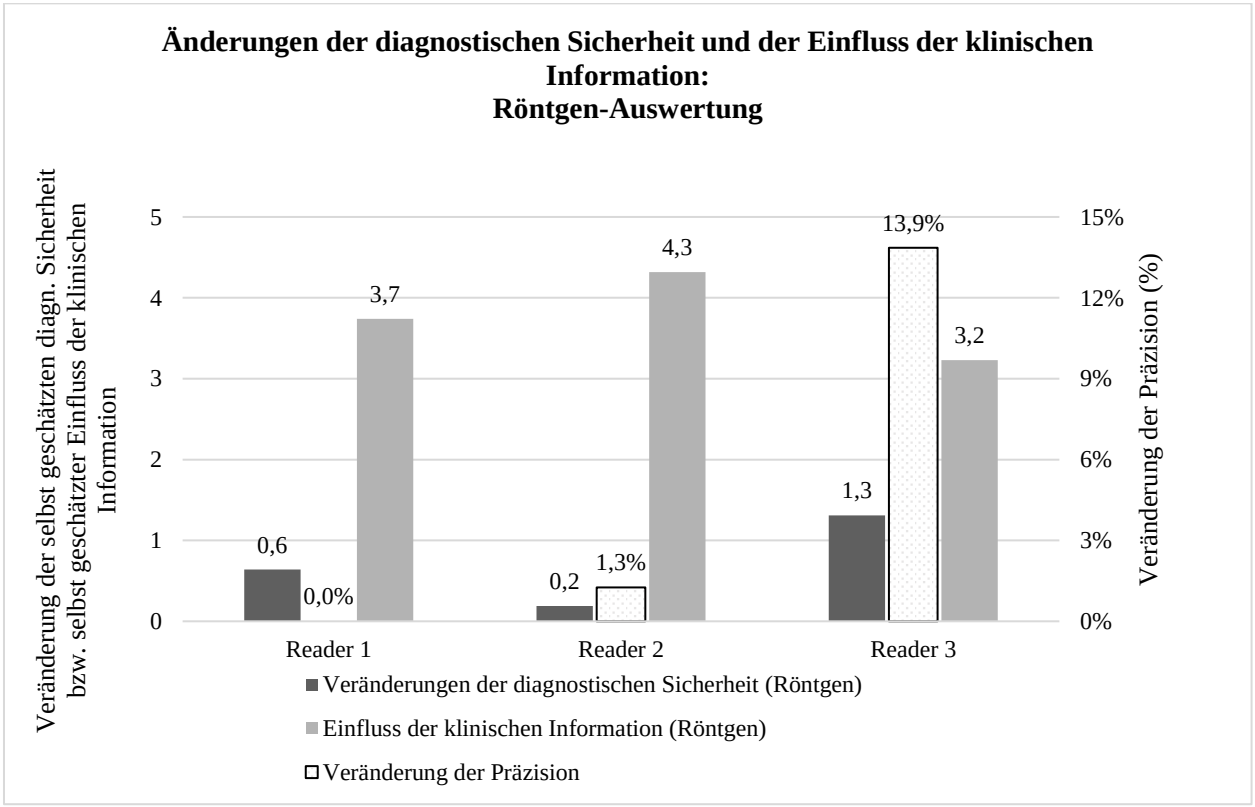


Abbildung 10 Diagnostische Sicherheit und Einfluss der klinischen Information pro Reader (Röntgen)
Veränderungen der subjektiv geschätzten diagnostischen Sicherheit sowie der Präzision (Differenz aus der Auswertung ohne vs. mit klinischer Information) im Vergleich zum subjektiv geschätzten Einfluss der klinischen Information auf die Befundung: Röntgen (Runden 1 und 2), separate Darstellung der drei Reader. Präzision: Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel“ bzw. „nicht mit der Diagnose axSpA kompatibel“ bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); axSpA: axiale Spondyloarthritis.

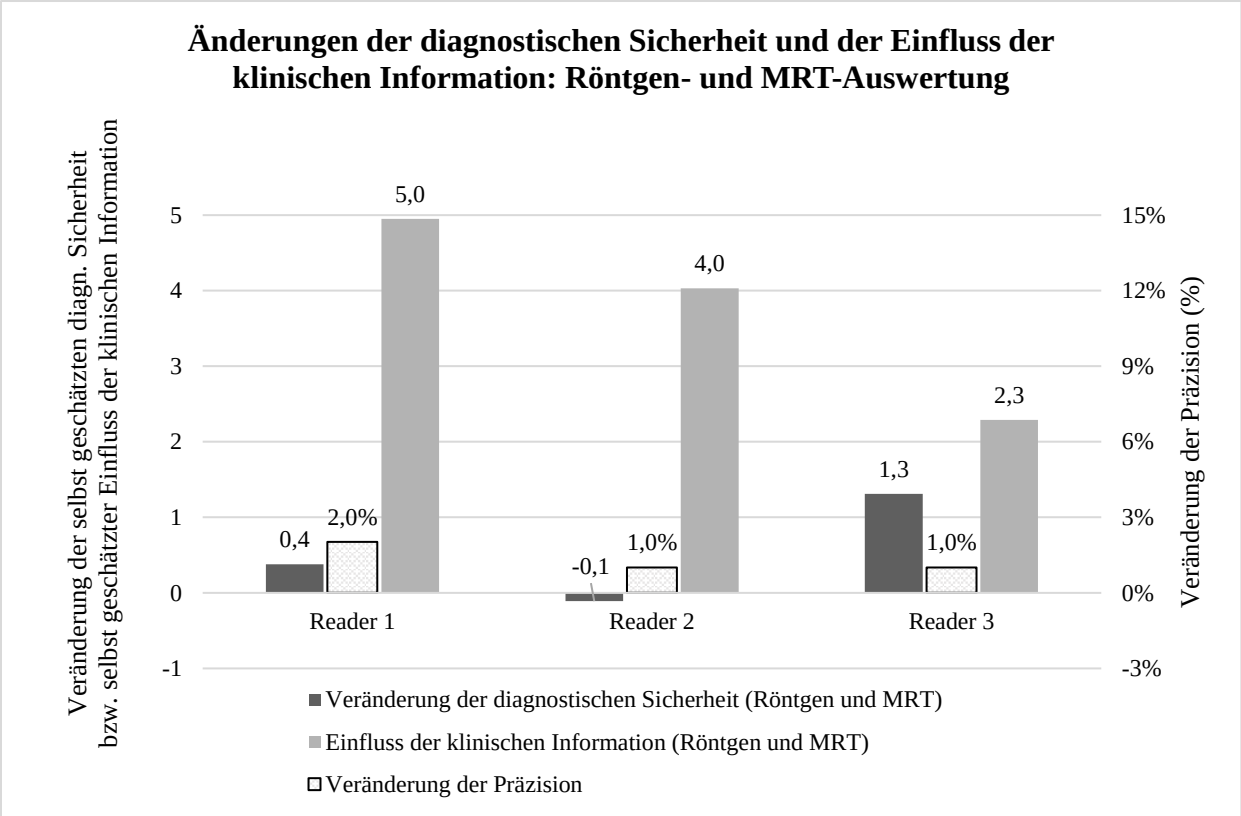


Abbildung 11 Diagnostische Sicherheit und Einfluss der klinischen Information pro Reader (Röntgen und MRT)
Veränderung der subjektiv geschätzten diagnostischen Sicherheit sowie der Präzision (Differenz aus der Auswertung ohne vs. mit klinischer Information) im Vergleich zum subjektiv geschätzten Einfluss der klinischen Information auf die Befundung: Röntgen und MRT (Runden 3 und 4), separate Darstellung der drei Reader. Präzision: Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel“ bzw. „nicht mit der Diagnose axSpA kompatibel“ bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); axSpA: axiale Spondyloarthritis.

3.8 Analyse der Befundung in Abhängigkeit von der subjektiv geschätzten diagnostischen Sicherheit

Bei der nach Diagnosesicherheit kategorisierten Analyse der Röntgen-Befundung zeigt sich, dass mit zunehmender selbst geschätzter Sicherheit die Präzision steigt, unabhängig von der klinischen Information. In allen drei „Sicherheitsgruppen“ zeigt sich eine Zunahme der Präzision durch das Vorliegen der klinischen Information im Vergleich zur Befundung ohne klinische Information. Am stärksten ausgeprägt ist diese Zunahme bei den „unsicheren“ Fällen (selbst geschätzte diagnostische Sicherheit von 0 bis 3) – hier steigt die diagnostische Genauigkeit um circa 20%, während sie in den Fällen mit mittlerer und hoher selbst geschätzter Sicherheit um 4 – 6% zunimmt (s. Abbildung 12).

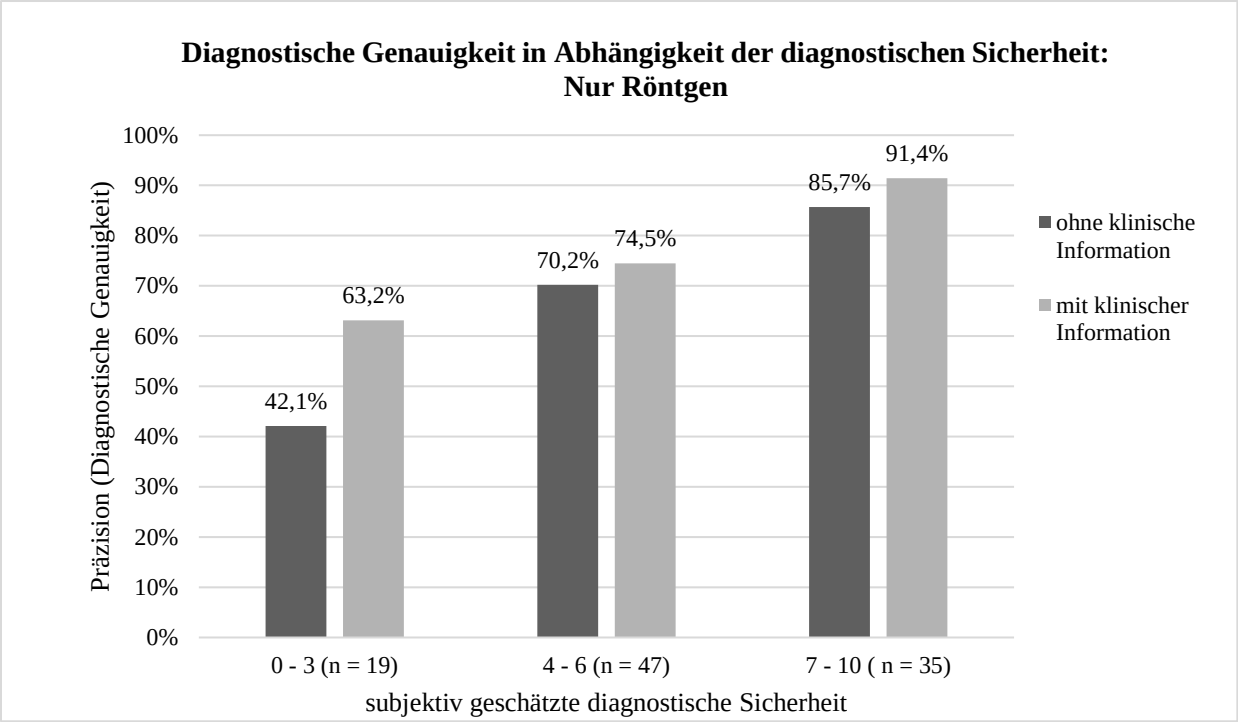


Abbildung 12 Diagnostische Genauigkeit in Abhängigkeit der Sicherheit: Röntgen
Diagnostische Genauigkeit der Röntgenauswertung (Runden 1 und 2) in Abhängigkeit der durch die Radiolo*innen geschätzten diagnostischen Sicherheit. 0-3: geringe, 4-6: moderarte, 7-10 hohe diagnostische Sicherheit. Aggregierte Daten der drei Radiolog*innen. Präzision: Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel“ bzw. „nicht mit der Diagnose axSpA kompatibel“ bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); axSpA: axiale Spondyloarthritis.

Darüber hinaus zeigt sich in Tabelle 16, dass der selbst geschätzte Einfluss der klinischen Information auf die Befundung der Röntgenaufnahmen bei gering geschätzter Sicherheit am größten ist (5.1) und umgekehrt bei zunehmender Sicherheit sinkt (2.5). Die Sensitivität der Röntgenbefundung steigt sowohl mit zunehmender selbst geschätzter diagnostischer Sicherheit (von 23% auf 71%) als auch bei Vorliegen der klinischen Information. Bei „unsicheren“ Fällen ist dieser Effekt am größten, hier zeigt sich nach Kenntnisnahme der klinischen Information ein Zuwachs der Sensitivität von > 20% im Vergleich zur Befundung ohne klinische Information. Die Spezifität beträgt bei vorliegender klinischer Information in allen „Sicherheitsstufen“ 100%, während sie ohne klinische Information bei „unsicheren“ Fällen etwas niedriger liegt als bei Fällen, die mit einer diagnostischen Sicherheit von ≥ 4 eingeschätzt wurden (83% vs. 95%). Die selbst geschätzte diagnostische Sicherheit steigt bei Vorliegen der klinischen Information in der Gruppe mit initial niedrigster Sicherheit am stärksten (+1,6) an.

Bei der detaillierten Analyse der drei Radiolog*innen zeigt sich eine deutliche Variabilität in der Einschätzung der diagnostischen Sicherheit (Anzahl der Fälle, die mit einer niedrigen

Sicherheit eingeschätzt wurden: Bereich 18 – 35; Anzahl der Fälle, die mit hoher Sicherheit eingeschätzt wurden: Bereich 13 - 56). Es zeigt sich auch hier die Tendenz, dass die Präzision sowohl mit zunehmender selbst geschätzter diagnostischer Sicherheit als auch nach Kenntnisnahme der klinischen Information steigt (s. Tabelle 18).

Diagnostische Sicherheit[1] (ohne klinische Information)	**n**	**Klinische Information:**	**Präzision [2]**	**Diagn. Sicherheit [1,3]**	**Einfluss d. klinischen Information[4,3]**	**SE**	**SP**	**PPV**	**NPV**
Niedrig:	19	Nein	42.1 %	2.7		0.23	0.83	0.75	0.33
0 – 3		*Ja*	*63.2 %*	*4.3*	*5.1*	*0.46*	*1*	*1*	*0.46*
Moderat:	47	Nein	70.2 %	4.7		0.54	0.95	0.94	0.58
4 – 6		*Ja*	*74.5 %*	*5.9*	*4.1*	*0.57*	*1*	*1*	*0.61*
Hoch:	35	Nein	85.7 %	7.4		0.71	0.95	0.91	0.83
7 – 10		*Ja*	*91.4 %*	*6.9*	*2.5*	*0.79*	*1*	*1*	*0.88*

Tabelle 16 Röntgen-Befundung in Abhängigkeit der subjektiv geschätzten diagnostischen Sicherheit

[1]Selbst geschätzt durch Radiolog*innen (0 = min. Sicherheit, 10 = max. Sicherheit); [2]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); [3]Mittelwerte; [4]Selbst geschätzter Einfluss der klinischen Information auf die Befundung (0 = min. Einfluss; 10 = max. Einfluss). Abk.: SE: Sensitivität; SP: Spezifität; PPV: Positiv prädiktiver Wert; NPV: negativ prädiktiver Wert; axSpA: axiale Spondyloarthritis. Aggregierte Werte der drei Radiolog*innen.

Die Ergebnisse der Röntgen- und MRT-Befundung (Runden 3 und 4), dargestellt nach selbst geschätzter diagnostischer Sicherheit, sind in Abbildung 13 zu sehen. Hier fällt zunächst auf, dass die Anzahl der Fälle, die als „unsicher" eingestuft wurden deutlich geringer ist als dies bei der alleinigen Röntgen-Auswertung der Fall ist (4 vs. 19). Entsprechend wurden fast doppelt so viele Befunde mit einer hohen diagnostischen Sicherheit bewertet (60 vs. 35). Auch hier steigt, korrelierend mit der Rötgenbefundung, die diagnostische Genauigkeit (Präzision) weitestgehend mit zunehmender Sicherheit an. Wurden die Röntgen- und MRT-Bilder mit der klinischen Information zusammen ausgewertet, so beobachtet man in den verschiedenen „Sicherheitsgruppen" sowohl Ab- als auch Zunahmen der Präzision, die jedoch mit Ausnahme der wenigen „unsicheren" Fälle unter 4% betragen.

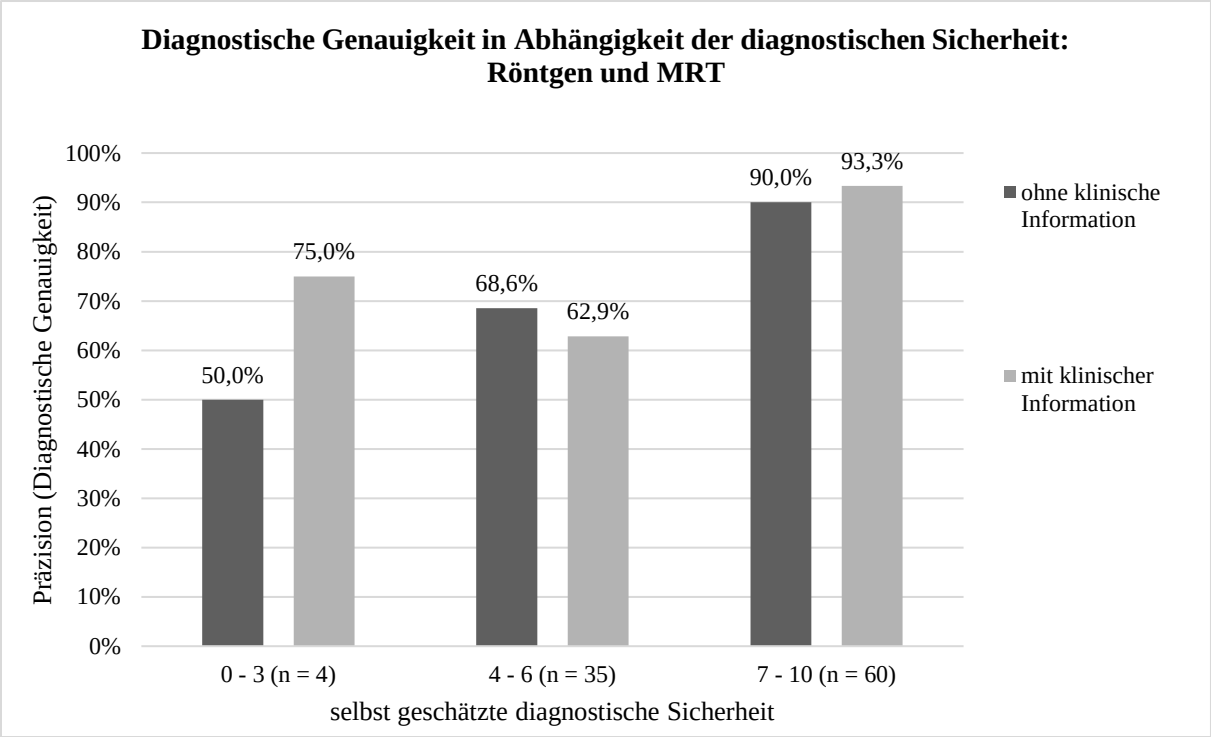


Abbildung 13 Diagnostische Genauigkeit in Abhängigkeit der Sicherheit: Röntgen und MRT
Diagnostische Genauigkeit (Präzision) der Röntgen- und MRT-Auswertung (Runden 3 und 4) in Abhängigkeit der durch die Radiolo*innen geschätzten diagnostischen Sicherheit. 0-3: geringe, 4-6: moderate, 7-10 hohe diagnostische Sicherheit. Aggregierte Daten der drei Radiolog*innen. Präzision: Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); axSpA: axiale Spondyloarthritis.

Wie auch bei der alleinigen Röntgenauswertung zeigt sich bei den Ergebnissen der Röntgen- und MRT-Befundung ein Anstieg der selbst geschätzten diagnostischen Sicherheit nach Kenntnisnahme der klinischen Information. Dieser ist bei den Fällen mit niedriger geschätzter Sicherheit stärker ausgeprägt (+1.9) als bei den Fällen mit hoher Sicherheit (+0.2). Der von den Radiolog*innen geschätzte Einfluss der klinischen Information auf die Befundung verhält sich wiederum umgekehrt zur Einschätzung der diagnostischen Sicherheit.

Die Sensitivität steigt ebenfalls mit Zunahme der selbst berichteten diagnostischen Sicherheit (ohne klinische Information), während das Vorliegen der klinischen Information unterschiedliche Auswirkungen aufweist: In der Gruppe mit niedriger Sicherheit steigt sie deutlich, bei moderater Sicherheit sinkt sie geringfügig und in der Gruppe mit hoher Sicherheit zeigt sich keine Änderung der Sensitivität. In den Fällen, die von den Radiolog*innen mit einer moderaten oder hohen Sicherheit eingeschätzt wurden, steigt die Spezifität sowohl mit zunehmender Sicherheit des Readers als auch bei Vorliegen der klinischen Information (s. Tabelle 17).

Diagnostische Sicherheit[1] (ohne klinische Information)	**n**	**Klinische Information**	**Präzision[2]**	**Diagn. Sicher-heit[3]**	**Einfluss der klinischen Information[4,3]**	**SE**	**SP**	**PPV**	**NPV**
Niedrig:	4	Nein	50.0 %	2.9		0.33	1	1	0.33
0 – 3		*Ja*	*75.0 %*	*4.8*	*4.8*	*0.67*	*1*	*1*	*0.5*
Moderat:	35	Nein	68.6 %	5.3		0.57	0.86	0.86	0.57
4 – 6		*Ja*	*62.9 %*	*6.2*	*4.7*	*0.43*	*0.93*	*0.9*	*0.52*
Hoch:	60	Nein	90.0 %	7.7		0.88	0.92	0.94	0.86
7 – 10		*Ja*	*93.3 %*	*7.9*	*3.1*	*0.88*	*1*	*1*	*0.87*

Tabelle 17 Röntgen- und MRT-Befundung in Abhängigkeit der subjektiv geschätzten diagnostischen Sicherheit

[1]Selbst geschätzt durch Radiolog*innen (0 = min. Sicherheit, 10 = max. Sicherheit); [2]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen (Goldstandard); [3]Mittelwerte; [4]Selbst geschätzter Einfluss der klinischen Information auf die Befundung (0 = min. Einfluss; 10 = max. Einfluss). Abk.: axSpA: axiale Spondyloarthritis; SE: Sensitivität; SP: Spezifität; PPV: Positiv prädiktiver Wert; NPV: negativ prädiktiver Wert. Aggregierte Werte der drei Radiolog*innen.

Bei der Einzelauswertung der Reader zeigt sich ähnlich wie bei der Röntgen-Auswertung eine Variabilität in der Beurteilung der Fälle, die mit niedriger (Bereich (n) 7 – 16) bzw. hoher (Bereich (n) 44 – 72) diagnostischer Sicherheit eingeschätzt werden (s. Tabelle 19).

Diagnostische Sicherheit[1]		Klin. Info.:	n	Präzision[2]	Diagn. Sicher-heit[3]	Einfluss d. klin. Info.[3,4]	SE	SP	PPV	NPV
R1	Niedrig: 0 – 3	Nein	18	72.2 %	2.3		0.73	0.71	0.8	0.63
		Ja		*77.8 %*	*5.0*	*4.8*	*0.64*	*1*	*1*	*0.64*
	Moderat: 4 – 6	Nein	57	71.9 %	5.1		0.52	0.96	0.94	0.63
		Ja		*70.2 %*	*5.7*	*3.7*	*0.45*	*1*	*1*	*0.61*
	Hoch: 7 – 10	Nein	26	88.5 %	7.7		0.85	0.92	0.92	0.86
		Ja		*88.5 %*	*7.2*	*3.2*	*0.77*	*1*	*1*	*0.81*
R2	Niedrig: 0 – 3	Nein	27	40.7 %	2.1		0.29	0.54	0.4	0.41
		Ja		*48.1 %*	*3.6*	*5.2*	*0.21*	*0.77*	*0.5*	*0.48*
	Moderat: 4 – 6	Nein	61	67.2 %	5.1		0.53	0.81	0.73	0.64
		Ja		*65.6 %*	*4.9*	*4.3*	*0.5*	*0.81*	*0.71*	*0.63*
	Hoch: 7 – 10	Nein	13	92.3 %	7.6		0.91	1	1	0.67
		Ja		*92.3 %*	*6.8*	*2.4*	*0.91*	*1*	*1*	*0.67*
R 3	Niedrig: 0 – 3	Nein	35	45.7 %	1.8		0.25	0.91	0.86	0.36
		Ja		*68.6 %*	*6.7*	*5.0*	*0.58*	*0.91*	*0.93*	*0.5*
	Moderat: 4 – 6	Nein	10	60.0 %	4.9		0.33	1	1	0.5
		Ja		*70.0 %*	*6.4*	*3.9*	*0.67*	*0.75*	*0.8*	*0.6*
	Hoch: 7 – 10	Nein	56	75.0 %	8.6		0.52	0.94	0.87	0.71
		Ja		*83.9 %*	*7.6*	*2.0*	*0.64*	*1*	*1*	*0.78*

Tabelle 18 Röntgen-Befundung in Abhängigkeit der subjektiv geschätzten diagnostischen Sicherheit separat pro Reader

[1]Selbst geschätzt durch Radiolog*innen (0 = min. Sicherheit, 10 = max. Sicherheit); [2]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen; [3]Mittelwerte; [4]Selbst geschätzter Einfluss der klinischen Information auf die Befundung (0 = min. Einfluss; 10 = max. Einfluss); Abk.: SE: Sensitivität; SP: Spezifität; PPV: Positiv prädiktiver Wert; NPV: negativ prädiktiver Wert. axSpA: axiale Spondyloarthritis; R: Reader.

Diagnostische Sicherheit[1]		Klin. Info.:	n	Präzision[2]	Diagn. Sicher-heit[3]	Einfluss d. klin. Info. [3,4]	SE	SP	PPV	NPV
R1	Niedrig: 0 – 3	Ohne	7	71.4 %	2.7		0.6	1	1	0.50
		mit		*57.1 %*	*3.9*	*6.9*	*0.4*	*1*	*1*	*0.40*
	Moderat: 4 – 6	Ohne	30	60.0 %	5.6		0.38	0.86	0.75	0.55
		Mit		*66.7 %*	*6.8*	*5.3*	*0.38*	*1*	*1*	*0.58*
	Hoch: 7 – 10	Ohne	62	90.3 %	7.4		0.89	0.92	0.94	0.85
		Mit		*91.9 %*	*7.3*	*4.6*	*0.92*	*0.92*	*0.94*	*0.89*
R2	Niedrig: 0 – 3	Ohne	7	57.1 %	1.7		0.5	0.6	0.33	0.75
		Mit		*42.9 %*	*3.7*	*4.4*	*0.0*	*0.6*	*0.0*	*0.6*
	Moderat: 4 – 6	Ohne	48	62.5 %	4.9		0.55	0.69	0.6	0.64
		Mit		*68.8 %*	*5.2*	*4.3*	*0.5*	*0.85*	*0.73*	*0.67*
	Hoch: 7 – 10	Ohne	44	95.5 %	7.9		0.97	0.9	0.97	0.9
		Mit		*93.2 %*	*7.1*	*3.7*	*0.94*	*0.9*	*0.97*	*0.82*
R 3	Niedrig: 0 – 3	Ohne	16	62.5 %	3.0		0.58	0.75	0.88	0.38
		Mit		*56.3 %*	*6.6*	*4.1*	*0.42*	*1*	*1*	*0.36*
	Moderat: 4 – 6	Ohne	11	63.6 %	5.1		0.56	1	1	0.33
		Mit		*63.6 %*	*7.5*	*4.2*	*0.56*	*1*	*1*	*0.33*
	Hoch: 7 – 10	Ohne	72	90.3 %	9.0		0.81	1	1	0.83
		Mit		*93.1 %*	*9.4*	*1.6*	*0.87*	*1*	*1*	*0.88*

Tabelle 19 Röntgen- und MRT-Befundung in Abhängigkeit der subjektiv geschätzten diagnostischen Sicherheit separat pro Reader

[1]Selbst geschätzt durch Radiolog*innen (0 = min. Sicherheit, 10 = max. Sicherheit); [2]Übereinstimmung der durch die Radiolog*innen (Reader) als „mit der Diagnose axSpA kompatibel" bzw. „nicht mit der Diagnose axSpA kompatibel" bewerteten Fälle mit der klinischen Diagnose durch Rheumatolog*innen; [3]Mittelwerte; [4]Selbst geschätzter Einfluss der klinischen Information auf die Befundung (0 = min. Einfluss; 10 = max. Einfluss); Abk.: SE: Sensitivität; SP: Spezifität; PPV: Positiv prädiktiver Wert; NPV: negativ prädiktiver Wert. axSpA: axiale Spondyloarthritis; R: Reader.

4 Diskussion:

Die oben vorgelegten Ergebnisse zeigen, dass die zur Verfügung gestellten klinischen Informationen insgesamt einen positiven Einfluss auf die radiologische Befundung haben. Dabei zeigt sich, vor allem hinsichtlich der Röntgenbefunde, deutliche Unterschiede im Vergleich vor und nach Erhalt der klinischen Informationen. Es gibt aber auch Unterschiede sowohl zwischen der alleinigen Röntgenauswertung und der kombinierten Röntgen- und MRT-Auswertung als auch zwischen den Ergebnissen der einzelnen Reader.

Die Beantwortung der beiden Hauptfragen ist für die alleinige Röntgen-Auswertung klarer: die festgestellte Präzision in Bezug auf die klinische Diagnose axSpA beträgt vor der klinischen Information 70,3% und danach 78.2% - dieser Unterschied ist statistisch signifikant (p=0,008). Bei der Kombination von Röntgen und MRT liegt die Präzision ohne klinische Information bei 80.8% und mit bei 81.8%. Dieser Unterschied ist statistisch nicht signifikant. Daraus lässt sich schließen, dass die klinische Information bei der Röntgen-Befundung mehr Vorteile bringt als bei der MRT-Befundung, was wahrscheinlich auf die niedrigere diagnostische Zuverlässigkeit der Röntgenmethodik zurückzuführen ist.

Grundsätzlich ist die MRT-Befundung aber präziser, was zum Teil dadurch zu erklären ist, dass insgesamt nur 59,6% der axSpA-Patienten definitive Röntgenveränderungen in den SIG (mNYC+) hatten, während mittels MRT bei 67,2% strukturelle Veränderungen gesehen wurden und bei 65,5% entzündliche Veränderungen.

Andererseits könnte auch die simultane Sichtbarkeit von entzündlichen und strukturellen Veränderungen im MRT (+Röntgen), die bei den meisten axSpA-Patient*innen gegeben ist, zur höheren Präzision der Radiolog*innen gegenüber der alleinigen Röntgen-Befundung beigetragen haben. Die Sichtbarkeit beider Veränderungen im MRT wurden in einer kürzlich publizierten Arbeit als besonders Diagnose-weisend bei axSpA-Patient*innen identifiziert (29).

Entsprechend liegt die Sensitivität für die klinische Diagnose axSpA beim Röntgen ohne klinische Information nur bei 51% und mit bei 60%. Die Spezifität steigt aber von 93% auf 100% an, und der PPV liegt letztlich sogar bei 100%. Dabei steigt auch der Kappa-Wert für die Interrater-Reliabilität von einer moderaten ($\kappa = 0.56$) auf eine gute (substanzielle) Übereinstimmung ($\kappa = 0.66$) an.

Im Vergleich dazu liegt die Sensitivität für die klinische Diagnose axSpA bei der Röntgen- und MRT-Auswertung ohne klinische Information bei 74% und mit bei 71%. Die Spezifität steigt

aber von 90% auf 98% an, und der PPV liegt letztlich bei 98%. Dabei zeigt der Kappa-Wert zwar eine grundlegend höhere Übereinstimmung der Reader als bei der alleinigen Röntgenbefundung, durch die klinische Information jedoch kaum eine Veränderung (von κ= 0,77 auf κ= 0,75). Beide Werte entsprechen einer guten (substanziellen) Übereinstimmung.

Insgesamt liegt nahe, dass die klinischen Informationen, wenn hinweisend auf axSpA, die fraglicheren Befunde eher zu positiven im Sinne der axSpA gemacht haben und umgekehrt, bei fehlenden weiteren Hinweisen auf axSpA, der Befund der Radiolog*innen nicht im Sinne einer axSpA gewertet wurde. Dafür spricht auch die Auswertung der diagnostischen Genauigkeit in Abhängigkeit von der subjektiv empfundenen diagnostischen Sicherheit (vgl. Abbildung 12 und Tabelle 16).

Hierbei zeigt sich die größte Zunahme der diagnostischen Genauigkeit, also der Übereinstimmung mit dem Goldstandard, der Sensitivität und der selbst geschätzten Sicherheit bei den Fällen mit der initial niedrigsten diagnostischen Sicherheit, d.h. bei den Fällen, die ohne klinische Information als eher „unsicher“ befundet wurden.

Bei der genaueren Betrachtung der radiologischen Befunde zeigt sich, dass bei der Röntgenauswertung mit der Kenntnis der klinischen Information deutlich häufiger definitive strukturelle Veränderungen im Sinne der mNYC festgestellt wurden (s. Abschnitt 3.4). Zwar ist dieser Unterschied nicht signifikant, dennoch könnte dies erklären, dass bzw. warum mittels klinischer Information mehr Strukturveränderungen im Röntgen erkannt wurden, was die Steigerung der Sensitivität erklärt und gleichzeitig dazu führt, dass weniger bzw. in dieser Analyse sogar kein Röntgenbild fälschlicherweise als positiv im Sinne von definitiven röntgenologischen Veränderungen gewertet wurde. Letzteres drückt sich durch den Anstieg der Spezifität aus.

Dem gegenüber sind die Veränderungen der Befundung bei den im MRT sichtbaren SpA-typischen Veränderungen durch die Kenntnis der klinischen Information in umgekehrter Richtung zu verzeichnen: hier nimmt der Anteil der von den Radiolog*innen gesehenen aktiv entzündlichen und strukturellen Veränderungen ab, wobei der Unterschied der letzteren sogar signifikant ist (s. Tabelle 9). Vor dem Hintergrund der dabei deutlich gestiegenen Spezifität lässt sich vermuten, dass die Radiolog*innen durch die Kenntnisnahme der klinischen Information kritischer in der Befundung SpA-typischer Veränderungen werden und somit weniger falsch positive Befunde entstehen. Gleichzeitig zeigt sich aber durch die leicht gesunkene Sensitivität, dass ebenso weniger axSpA-Patient*innen erfasst werden.

Die Übereinstimmung der Radiolog*innen bezüglich der SpA-typischen MRT-Veränderungen liegt in einem guten (substanziellen) Bereich und ist vergleichbar mit den Ergebnissen anderer Studien (35, 45).

Der für diese Studie festgelegte Goldstandard ist die klinische Diagnose der Rheumatolog*innen. Das bedeutet, dass die Diagnose auch durchaus bei Patienten gestellt werden konnte, die gar keine Veränderungen in der Bildgebung hatten. Nach entsprechender Überprüfung dieses Sachverhalts zeigt sich, dass in der Tat bei ca. 30% der Patient*innen mit der klinischen Diagnose axSpA keine pathologische Bildgebung hinsichtlich Röntgen und MRT vorliegen. Das erklärt dann auch die mit 70% relativ niedrige Sensitivität der radiologischen Befunde für die Diagnose axSpA. In dem Zusammenhang ist noch von Interesse, dass sich bei 62% der Patient*innen mit axSpA in der MRT beides, Zeichen von Entzündung und Strukturveränderungen, nachweisen ließen. Die Radiolog*innen haben auch überprüft, ob in den Röntgenbildern die mNYC erfüllt sind – dies ist bei 60% der axSpA der Fall. Unter Berücksichtigung der mit 6,7 Jahren doch schon etwas längeren Krankheitsdauer der axSpA-Patient*innen stellt dies einen durchaus zu erwartenden Befund dar. Auch die HLA B27-Häufigkeit liegt in dieser Gruppe mit 84,5% im erwarteten Bereich (25). Die „mNYC-positiven“ Patient*innen haben auch insgesamt mehr Entzündung und mehr Strukturveränderungen als die „mNYC-negativen“ - das entspricht veröffentlichten Daten aus anderen Studien (66).

Allerdings bestätigen sich die in der Fallzahlberechnung angenommenen Teilmengen nur zum Teil. Wie oben beschrieben wurde vorab eine Diskrepanz zwischen dem radiologischen Befund und der rheumatologischen Diagnose vor Zurverfügungstellung der klinischen Information von 20% angenommen. Diese sollte sich danach auf 5% verringern. Formal richtig ist, dass bei der Röntgenauswertung ohne klinische Information 70% und mit 78% Übereinstimmung und bei der Röntgen- und MRT-Auswertung ohne klinischer Information 81% und mit 82% Übereinstimmung mit der klinischen Diagnose vorliegen. Da jedoch ein nicht zu vernachlässigender Teil der axSpA-Patient*innen gar keine Veränderungen in der Bildgebung aufweisen, müssen diese Zahlen unter diesem Aspekt betrachtet werden.

4.1 Diagnostische Sicherheit und Einfluss der klinischen Information:

Darüber ist festzustellen, dass sowohl in der Röntgendiagnostik als auch in der Befundung des gesamten zur Verfügung stehenden Bildmaterials (Röntgen und MRT) die klinische

Information in jeweils circa der Hälfte der Fälle zu einer Zunahme der gefühlten Sicherheit der befundenden Person führt und nur bei einem Viertel (Röntgen) bzw. sogar bei nur 7% (Röntgen und MRT) zu einer Verunsicherung (s. Tabelle 14).

Der Einfluss der klinischen Information wurde im Vergleich zwischen alleiniger Röntgen- vs. Röntgen- und MRT-Befundung ähnlich, aber relativ gering (Mittelwert < 4 auf einer 0-10 Skala) eingeschätzt. Die etwas größere Streuung beim Röntgen könnte darauf hindeuten, dass es mehr unklare Fälle gab. Bei der kombinierten Röntgen- und MRT-Befundung liegen weniger Abweichungen vor.

Unabhängig von der Bildmodalität zeigt sich, dass der Einfluss der klinischen Information bei Fällen höher eingeschätzt wird, in denen sich die Radiolog*innen initial unsicher waren. Mit der Erkenntnis, dass in der Mehrzahl der Fälle die selbstgeschätzte Sicherheit der Befunder*innen mit der klinischen Information steigt oder gleichbleibt, lässt sich also festhalten, dass dieser Einfluss eine eher positive Wirkung auf die Reader hat. Dies wird auch daran deutlich, dass gerade bei unsicheren Fällen die Zunahme der selbst geschätzten Sicherheit der Reader am größten ausgeprägt ist (s. Abbildung 12 und Abbildung 13).

Abbildung 10 und Abbildung 11 veranschaulichen, dass die Einschätzung des Einflusses der klinischen Information eine deutliche Variabilität aufweist. Die Reader 1 und 2 schätzen den Einfluss der klinischen Information beim Röntgen und MRT höher ein als Reader 3. Nach klinischer Information ist aber keine größere Veränderung der Präzision und Sicherheit bei ihnen zu konstatieren. Dagegen schätzt Reader 3 (die Person mit der größten Expertise) den Einfluss der klinischen Information sowohl bei der alleinigen Röntgenbefundung als auch bei der kombinierten Röntgen- und MRT-Auswertung am geringsten ein – seine/ihre diagnostische Sicherheit und Präzision nimmt jedoch bei vorliegender klinischer Information in der Röntgenauswertung deutlich zu. Es zeigt sich also eine Diskrepanz zwischen wahrgenommenem Einfluss der klinischen Information und der resultierenden Veränderung der Befundung, wobei diese Effekte ebenso abhängig von der Expertise der befundenden Person zu sein scheinen.

4.2 Variabilität der Reader:

In der Tat gibt es zu verzeichnende Diskrepanzen bei den drei unterschiedlich erfahrenen Readern. Bei der Röntgenbefundung liegt die Präzision ohne klinische Information bei 63% - 76% und mit bei 64 – 77%. Noch größere Unterschiede finden sich bei der Sensitivität: ohne

klinische Information 38% - 64% und mit 51% - 62%, während die Spezifität sich zwischen 74% und 94% sowie 81% und 100% bewegt. Hierbei liegt die Sensitivität bei einem Reader ohne klinische Information bei nur 38%. Mit klinischer Information zeigt sich ein erheblicher Anstieg auf 62% – dies findet sich in einer erheblichen Verbesserung der Präzision von 63% auf 77% wieder.

Bei der kombinierten Röntgen- und MRT-Befundung liegt die Variabilität der Präzision ohne klinische Information bei 77% - 83% und mit bei 79 - 84%. Die Unterschiede in der Sensitivität sind deutlich geringer als bei der alleinigen Röntgenauswertung ausgeprägt: ohne klinische Information 72% - 79% und mit 72% - 74%, während die Spezifität sich zwischen 73% und 98% sowie 83% und 100% bewegt.

Zwischen den Readern finden sich also zum Teil erhebliche Diskrepanzen – sowohl hinsichtlich der Einschätzung axSpA-kompatibel vs. nicht axSpA-kompatibel als auch im Hinblick auf Sensitivität und Spezifität. Grundsätzlich verändern sich die Werte bei der alleinigen Röntgenbefundung positiv, während dies bei der Röntgen- und MRT-Auswertung zum Teil gegensinnig ist. Hier ist eine Zunahme der Spezifität von einem Verlust an Sensitivität begleitet.

Die diagnostische Sicherheit der Radiolog*innen ist bei der Kombination von Röntgen und MRT höher als beim alleinigen Röntgen. Die subjektive Einschätzung, wie groß der Einfluss klinischer Information war, zeigt, dass der Reader mit den geringsten Unterschieden bezüglich der Präzision vor vs. nach klinischer Information deren Einfluss am stärksten einschätzte. Auch hier zeigt sich eine nicht unerhebliche Variabilität.

4.3 Argumente pro und contra klinische Information

Für die Übermittlung umfangreicher klinischer Informationen an Radiolog*innen spricht, dass die diagnostische Genauigkeit der Befundung durch diese Erkenntnisse zunehmen kann (58, 59, 67-69). Darüber hinaus müssen für die radiologische Diagnostik wichtige Informationen zur Sicherheit der Patient*innen weitergegeben werden, z.B. zu Nierenfunktion und Kontrastmittelunverträglichkeit (70).

Andererseits befürchten manche Experten, dass die zusätzliche klinische Information die Radiolog*innen voreingenommen machen könnte und so eine objektive Begutachtung bei der Befundung und Interpretation der Bilder verhindert wird (71).

Von anderen Autoren ist ein Vorgehen empfohlen worden, bei dem zunächst die Befundung ohne weiteres Hintergrundwissen erfolgt, dann die klinischen Daten studiert werden, und daraufhin eine abschließende Interpretation durch die Radiolog*innen erfolgt – mit dem Ziel, die Inspektion und Wahrnehmung des gesamten Bildes sicherzustellen und damit die Rate falsch-positiver Befunde zu senken (59, 71). Dieses Vorgehen ist jedoch aufgrund des zusätzlichen Zeitaufwandes im normalen Versorgungsalltag wenig realistisch.

In einer systematischen Übersichtsarbeit, in der insgesamt 16 Studien berücksichtigt wurden (59), konnte gezeigt werden, dass die diagnostische Genauigkeit radiologischer Befunde zunimmt, wenn der befundenden Person klinische Informationen der Patient*innen vorliegen. Sieben der 16 eingeschlossenen Studien (44%) verwendeten dabei vordefinierte, strukturierte klinische Informationen. Dies wurde mit anderen Studien verglichen, in denen reale klinische Informationen angegeben wurden - hierbei schnitten erstere besser ab. D.h. eine deutlichere Verbesserung der diagnostischen Genauigkeit wurde durch die strukturierten klinischen Informationen im Vergleich mit den realen erreicht (86% Verbesserung versus 33%). Dieses Ergebnis spricht dafür, dass Radiolog*innen vor allem auf kondensierte Informationen (von hoher klinischer Relevanz) reagieren. Was nicht untersucht wurde, ist die Relevanz von Vorbefunden und früheren Bildern – auch dies kann den radiologischen Befund beeinflussen.

In einer weiteren Übersichtsarbeit aus 2020 wurde der Einfluss von klinischen Informationen auf die Diagnosegenauigkeit, die Diagnosesicherheit, die klinische Relevanz und die benötigte Zeit zur Befundung von CT-, MRT-, Röntgen und Ultraschallaufnahmen untersucht, wenn sie der befundenden Person vorab zu Verfügung gestellt wurden (58). Die Autoren kamen zu dem Schluss, dass die Zurverfügungstellung klinischer Informationen, verglichen mit den Befunden, die ohne Patient*innen-Information erhoben wurden, einen positiven Effekt auf fast alle untersuchten Parameter hatte – mit Ausnahme der Dauer der Befundung, hier wurde eine Zunahme beobachtet.

Zu ähnlichen Ergebnissen kam auch eine andere Untersuchung aus 2022, bei der die Befunde von CT-Aufnahmen von 64 onkologischen Patient*innen, jeweils mit und ohne klinische Informationen, von Radiolog*innen ausgewertet wurden (68). Dabei zeigte sich, dass sowohl die diagnostische Genauigkeit als auch die Sensitivität der Radiolog*innen hinsichtlich der Detektion von Primärtumoren und Metastasen durch die Kenntnis klinischer Informationen signifikant verbessert werden konnte.

In einer weiteren Studie zeigte sich ebenfalls ein Vorteil der Verfügbarkeit von klinischen Informationen für die radiologische Befundung (69). Dabei wurden CT-Aufnahmen von 410 Patient*innen mit akuten Bauchschmerzen retrospektiv von sechs Radiolog*innen mit und ohne klinische Informationen ausgewertet und die Diagnosegenauigkeit verglichen.

Als Radiolog*innen in Bezug auf die Qualität ihrer Kommunikation mit Klinikern befragt wurden, gehörte fehlende bzw. unzureichende klinische Informationen zu den größten berichteten Problemen (72). Das sind grundsätzlich vermeidbare Unzulänglichkeiten und Schwierigkeiten in der Qualität der Interpretation von Bildgebung.

Eine andere Befragung von Radiolog*innen zeigte, dass die Mehrheit einen deutlichen Vorteil in der adäquaten Kommunikation von klinischen Informationen sieht - sowohl durch Reduktion von Zufallsbefunden und Fehlern als auch durch die Verbesserung der Befunde (70).

Um die diagnostische Genauigkeit der Bildgebung bei axSpA zu erhöhen, sehen Radiolog*innen mit rheumatologischer Expertise einen Vorteil darin, wenn sie zur Beurteilung von Röntgen- und MRT-Aufnahmen bei Verdacht auf axSpA möglichst viel klinische Informationen erhalten (46). Den Autoren zufolge sollten Befunde wie Knochenmarködeme und strukturelle Läsionen in der MRT immer vor dem Hintergrund des klinischen Kontextes interpretiert werden, da einige Differenzialdiagnosen infrage kommen. Zu diesen zählen z.B. die Osteitis condensans ilii (73), die hauptsächlich bei postpartalen Frauen auftritt und mechanisch bedingte SIG-Veränderungen aufweist, die SIG-Arthrose, die häufiger bei älteren Patient*innen zu finden ist, traumatische oder stressbedingte Frakturen, gichtbedingte Veränderungen sowie auch die septische Sakroiliitis (44, 46).

Neben der verbesserten Beurteilbarkeit der vorliegenden Bildgebung, auch hinsichtlich möglicher Differenzialdiagnosen, ermöglicht die Kenntnis weiterer klinischer Parameter eine genauere und bessere Abschätzung der Wahrscheinlichkeit des Vorliegens der Diagnose einer axSpA. Zu diesen gehören neben dem normalerweise bekannten Alter und Geschlecht, die Art und Dauer der Symptome, der HLA-B27 Status und Entzündungswerte, die Familiengeschichte bezüglich SpA, SpA-assoziierte Manifestationen wie Uveitis, Psoriasis und CED, und auch die Wirksamkeit anti-entzündlicher Medikamente wie NSAR und Biologika.

4.4 Stellenwert der bildgebenden Methoden bei axSpA

In der Diagnostik der axSpA spielt die Bildgebung eine wichtige, oft sogar entscheidende Rolle. Falsch positive und falsch negative Befunde können die Diagnosestellung negativ beeinflussen,

indem sie dazu führen, dass Patient*innen fälschlicherweise die Diagnose axSpA erhalten und dann potenziell übertherapiert werden oder sogar trotz vorliegender Erkrankung nicht als axSpA identifiziert und damit nicht effektiv therapiert werden und möglicherweise einen Krankheitsprogress erleiden, der hätte verhindert werden können.

Die Diskussion über den Stellenwert der konventionellen Röntgenuntersuchung in der Diagnostik der axSpA hält in der rheumatologischen und radiologischen Fachgesellschaft seit Jahren an. Es gibt aber namhafte Autoren, die sich zuletzt bereits klar für die MRT als Standardmethode für die Bildgebung der SIG bei Verdacht auf axSpA positioniert haben (35), während die EULAR aktuell (seit 2015) in der diagnostischen Abklärung bei Verdacht auf axSpA immer noch die Röntgenuntersuchung als bildgebende Methode der ersten Wahl empfiehlt (38). Eine MRT sollte laut diesen Empfehlungen nur bei jungen Patient*innen mit kurzer Symptomdauer als primär eingesetzte Bildgebung vorrangig erwogen werden.

Neuere Untersuchungsergebnisse haben aber überzeugend nahegelegt, dass die MRT der Röntgenuntersuchung in Hinblick auf strukturelle Veränderungen der SIG bei axSpA in Sensitivität und Spezifität überlegen ist (35, 36, 42, 43). Dies stimmt mit den hier vorgelegten Ergebnissen überein.

Die Röntgenaufnahme der SIG zeigte im Gegensatz zur Kombination mit MRT auch in der hier vorgelegten Studie keinen zusätzlichen Nutzen. Die MRT könnte also in Zukunft – bei gegebener Verfügbarkeit – die konventionelle Röntgentechnik in der Diagnostik der SIG ablösen. Dies ist an der Wirbelsäule bekanntermaßen anders (74).

Der unzweifelhaft große Vorteil der MRT-Technik ist neben der präzisen und sensitiven Erkennung von strukturellen Veränderungen auch der mögliche Nachweis von entzündlichen Veränderungen, was mit der konventionellen Röntgentechnik nicht möglich ist.

Weitere MRT-Techniken wie die VIBE-Sequenz (volumetric interpolated breath-hold examination) zeigten bezüglich des Nachweises von Erosionen eine noch bessere Sensitivität (95% vs. 79%) bei gleicher Spezifität (jeweils 93%) im Vergleich zur T1-gewichteten Sequenz (75).

Auch die CT hat hinsichtlich Sensitivität (76% vs. 66%) und Spezifität (97% vs. 67%) gegenüber der konventionellen Röntgenuntersuchung der SIG deutliche Vorteile (76). Ein „low-dose“-CT wird aufgrund der vergleichbaren Strahlenexposition bei besserer Detektion struktureller Veränderungen und der höheren Interrater-Übereinstimmung der Radiolog*innen als mögliche bessere Alternative zur konventionellen Röntgentechnik diskutiert (76).

4.5 Was könnten diese Ergebnisse für die klinische Praxis bedeuten?

1. Die Ergebnisse dieser Arbeit sprechen dafür, den die SIG befundenden Radiolog*innen wichtige klinische Informationen hinsichtlich des möglichen Vorliegens einer axSpA zur Verfügung zu stellen und die mit der Anforderung verbundene Fragestellung klar zu formulieren.
2. Wenn bei axSpA-Patient*innen keine Veränderungen in der Bildgebung vorliegen, kann die Sensitivität des Befundes der Radiolog*innen in Bezug zur klinischen Diagnose nicht 100% erreichen – und liegt entsprechend dem Anteil normaler Befunde darunter.
3. In weniger häufigen Fällen gibt es bei Patient*innen mit axSpA primär eine Beteiligung der Wirbelsäule (mit entzündlichen und/oder strukturellen Veränderungen). Diese Bilder und/oder Informationen lagen den Radiolog*innen nicht vor und können die Befunde daher nicht beeinflusst haben.
4. Ein besonders hervorzuhebender Befund ist die von den Radiolog*innen zum Teil erreichte Spezifität von 98% - 100%. Das bedeutet, dass falsch positive Befunde in erfahrener Hand – bei vorliegender klinischer Information - weitgehend vermeidbar sein sollten.
5. Auch die Sensitivität ist hoch, wenn man nur die Fälle mit vorliegenden Veränderungen in der Bildgebung betrachtet.
6. Möglicherweise sollte, vor allem bei größerer Unsicherheit der Radiolog*innen hinsichtlich des Röntgenbefundes der SIG, mit den Kliniker*innen Rücksprache genommen werden, denn hier (bei größerer Unsicherheit) besteht offenbar das größte Verbesserungspotential.
7. Die Daten sprechen auch dafür, dass die Radiolog*innen die Kliniker*innen informieren sollten, wenn sie hinsichtlich des Befundes besonders sicher oder eben unsicher sind. Beispielsweise könnte eine Restunsicherheit des Befundes in Prozent angegeben werden.
8. Die Befunde der Radiolog*innen weisen ein nicht geringes Maß an Heterogenität und Variabilität auf. Das könnte an der unterschiedlichen Erfahrung der Beteiligten liegen.
9. Die hier erhobenen Daten beruhen auf der Performance von Radiolog*innen, die zumindest über eine gewisse Erfahrung in der Beurteilung von SIG bei axSpA verfügen. Bei weniger gut mit dem Thema vertrauten Radiolog*innen können die Ergebnisse anders aussehen. Daher erscheint es sinnvoll, die Expertise der befundenden Person transparent darzustellen.
10. Schlussendlich favorisieren unserer Ergebnisse die Weitergabe der klinischen Information an die Radiolog*innen, da sie zu einer Verbesserung der Präzision und der Spezifität der Bildbefundung und zu einer Vermeidung von Fehldiagnosen und Fehlklassifizierungen der axSpA führen kann.

5 Limitationen

Die vorliegende Studie hat einige Limitationen. Zunächst muss hier das retrospektive Studiendesign genannt werden, prospektive Studien sind aussagekräftiger. Die Tatsache, dass in die vorliegende Studie nur Patienten mit komplettem Bildmaterial (Röntgen- und MRT-Untersuchungen) eingeschlossen wurden, könnte einen Bias bewirken – möglicherweise wurden dadurch vermehrt schwerer betroffene Patient*innen eingeschlossen. Dafür spricht allerdings bei Betrachtung der demographischen Daten sonst nichts. Die Prävalenz von axSpA war in der vorliegenden Studie natürlich viel höher als in der Bevölkerung oder auch in einer allgemeinärztlichen oder rheumatologischen Praxis. Diese Anreicherung ist aber nötig, um valide Aussagen machen zu können.

In dieser Studie wurde nicht direkt untersucht, welchen Einfluss die klinische Information auf die Befundung der MRT-Bilder allein hat. Es ist jedoch anzunehmen, dass der Effekt sehr ähnlich zu dem auf die Auswertung von Röntgen- und MRT-Bildern gleichzeitig ist, wie es in dieser Studie der Fall war. Die Verfügbarkeit beider Bildmodalitäten bei axSpA-Patient*innen kommt dem klinischen Alltag dabei näher.

Methodisch kann nicht mit letzter Sicherheit ausgeschlossen werden, dass die Radiolog*innen sich beim ersten Durchgang (vor Erhalt der klinischen Information) Notizen gemacht haben, die ihnen dann beim zweiten Durchgang (nach Erhalt der klinischen Information) vorlagen. Ein Zugang zu weiteren bildgebenden Befunden derselben Patient*in (z.B. Aufnahmen der Wirbelsäule) war ebenfalls nicht völlig ausgeschlossen. Es wurde aber mehrfach versichert, dass dies nicht der Fall war bzw. dass diese Möglichkeit nicht genutzt wurde.

Die dargestellten Studienergebnisse beruhen zum Teil auf dem Prinzip der 2/3 Mehrheit, d.h. wenn zwei Radiolog*innen in ihrer Befundung übereinstimmten wurde dies als positiver Befund gewertet. Dieses wissenschaftsdemokratische Prinzip ist nicht per se fehlerfrei, stellt jedoch ein grundlegend akzeptiertes methodisches Vorgehen zur Identifikation positiver Befunde dar.

Inwieweit die Ergebnisse der Studie auf den klinischen Alltag übertragbar sind, muss offenbleiben. Tatsache ist aber, dass an dieser Studie vor allem Radiolog*innen mit mehr Erfahrung in der Bildgebung von axSpA beteiligt waren als im Normalfall zu erwarten ist. Die teilweise ausgeprägte Variabilität zwischen den Readern deutet darauf hin, dass die Befundung vor allem durch unerfahrenere Radiolog*innen bezüglich muskuloskelettaler Bildgebung anders ausfallen könnte.

Quellenverzeichnis

1. Diekhoff T, Eshed I, Giraudo C, Hermann KG, De Hooge M, Jans L, Jurik AG, Lambert RG, Machado PM, Maksymowych WP, Mallinson M, Marzo-Ortega H, Navarro-Compán V, Juhl Pedersen S, Østergaard M, Reijnierse M, Rudwaleit M, Sommerfleck F, Weber U, Baraliakos X, Poddubnyy D. OP0150 ASAS RECOMMENDATIONS FOR REQUESTING AND REPORTING IMAGING EXAMINATIONS IN PATIENTS WITH SUSPECTED AXIAL SPONDYLOARTHRITIS. Annals of the Rheumatic Diseases. 2022;81(Suppl 1):97.

2. Braun J, Sieper J. Ankylosing spondylitis. Lancet. 2007;369(9570):1379-90.

3. Kiltz U, Braun J, Dgrh, Becker A, Degam, Chenot JF, Dreimann M, Dwg, Hammel L, Dvmb, Heiligenhaus A, Dog, Hermann KG, Drg, Klett R, Dgmm, Krause D, Kreitner KF, Lange U, Dgpmr/Dgrw, Lauterbach A, Physio D, Mau W, Mossner R, Ddg, Oberschelp U, Dgooc, Philipp S, Pleyer U, Rudwaleit M, Schneider E, Schulte TL, Sieper J, Stallmach A, Dgim, Swoboda B, Dgooc/Dgorh, Winking M, Dgnc. [Long version on the S3 guidelines for axial spondyloarthritis including Bechterew's disease and early forms, Update 2019 : Evidence-based guidelines of the German Society for Rheumatology (DGRh) and participating medical scientific specialist societies and other organizations]. Z Rheumatol. 2019;78(Suppl 1):3-64.

4. Sieper J, Braun J. Clinician’s Manual on Axial Spondyloarthritis. 1 ed: Springer Healthcare Tarporley; 2014. XI, 96 p.

5. van Tubergen A, Heuft-Dorenbosch L, Schulpen G, Landewe R, Wijers R, van der Heijde D, van Engelshoven J, van der Linden S. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? Ann Rheum Dis. 2003;62(6):519-25.

6. Ramiro S, van der Heijde D, van Tubergen A, Stolwijk C, Dougados M, van den Bosch F, Landewe R. Higher disease activity leads to more structural damage in the spine in ankylosing spondylitis: 12-year longitudinal data from the OASIS cohort. Ann Rheum Dis. 2014;73(8):1455-61.

7. van Onna M, Jurik AG, van der Heijde D, van Tubergen A, Heuft-Dorenbosch L, Landewe R. HLA-B27 and gender independently determine the likelihood of a positive MRI of the sacroiliac joints in patients with early inflammatory back pain: a 2-year MRI follow-up study. Ann Rheum Dis. 2011;70(11):1981-5.

8. Navarro-Compán V, Sepriano A, El-Zorkany B, van der Heijde D. Axial spondyloarthritis. Annals of the Rheumatic Diseases. 2021;80(12):1511.

9. Stolwijk C, van Onna M, Boonen A, van Tubergen A. Global Prevalence of Spondyloarthritis: A Systematic Review and Meta-Regression Analysis. Arthritis Care Res (Hoboken). 2016;68(9):1320-31.

10. Brown MA. Human leucocyte antigen-B27 and ankylosing spondylitis. Intern Med J. 2007;37(11):739-40.

11. Braun J, Bollow M, Remlinger G, Eggens U, Rudwaleit M, Distler A, Sieper J. Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors. Arthritis Rheum. 1998;41(1):58-67.

12. Reveille JD, Weisman MH. The epidemiology of back pain, axial spondyloarthritis and HLA-B27 in the United States. Am J Med Sci. 2013;345(6):431-6.

13. Poddubnyy D, Sieper J. Similarities and differences between nonradiographic and radiographic axial spondyloarthritis: a clinical, epidemiological and therapeutic assessment. Curr Opin Rheumatol. 2014;26(4):377-83.

14. Gofton JP, Robinson HS, Trueman GE. Ankylosing spondylitis in a Canadian Indian population. Ann Rheum Dis. 1966;25(6):525-7.

15. Hanson A, Brown MA. Genetics and the Causes of Ankylosing Spondylitis. Rheum Dis Clin North Am. 2017;43(3):401-14.

16. Garrido-Mesa J, Brown MA. T cell Repertoire Profiling and the Mechanism by which HLA-B27 Causes Ankylosing Spondylitis. Curr Rheumatol Rep. 2022;24(12):398-410.

17. Xiong DK, Shi X, Han MM, Zhang XM, Wu NN, Sheng XY, Wang JN. The regulatory mechanism and potential application of IL-23 in autoimmune diseases. Front Pharmacol. 2022;13:982238.

18. Braun J, Brandt J, Listing J, Zink A, Alten R, Golder W, Gromnica-Ihle E, Kellner H, Krause A, Schneider M, Sorensen H, Zeidler H, Thriene W, Sieper J. Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. Lancet. 2002;359(9313):1187-93.

19. Baeten D, Sieper J, Braun J, Baraliakos X, Dougados M, Emery P, Deodhar A, Porter B, Martin R, Andersson M, Mpofu S, Richards HB, Group MS, Group MS. Secukinumab, an Interleukin-17A Inhibitor, in Ankylosing Spondylitis. N Engl J Med. 2015;373(26):2534-48.

20. Braun J, Kiltz U, Baraliakos X. Management of Axial Spondyloarthritis - Insights into Upadacitinib. Drug Des Devel Ther. 2022;16:3609-20.

21. Boel A, Lopez-Medina C, van der Heijde DMFM, van Gaalen FA. Age at onset in axial spondyloarthritis around the world: data from the Assessment in SpondyloArthritis international Society Peripheral Involvement in Spondyloarthritis study. Rheumatology. 2022;61(4):1468-75.

22. Sieper J, Poddubnyy D. Axial spondyloarthritis. The Lancet. 2017;390(10089):73-84.

23. de Winter JJ, van Mens LJ, van der Heijde D, Landewe R, Baeten DL. Prevalence of peripheral and extra-articular disease in ankylosing spondylitis versus non-radiographic axial spondyloarthritis: a meta-analysis. Arthritis Res Ther. 2016;18:196.

24. Rudwaleit M, Haibel H, Baraliakos X, Listing J, Marker-Hermann E, Zeidler H, Braun J, Sieper J. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum. 2009;60(3):717-27.

25. Braun J, Sieper J. Fifty years after the discovery of the association of HLA B27 with ankylosing spondylitis. RMD Open. 2023;9(3).

26. Braun J, Inman R. Clinical significance of inflammatory back pain for diagnosis and screening of patients with axial spondyloarthritis. Ann Rheum Dis. 2010;69(7):1264-8.

27. Sieper J, van der Heijde D, Landewe R, Brandt J, Burgos-Vagas R, Collantes-Estevez E, Dijkmans B, Dougados M, Khan MA, Leirisalo-Repo M, van der Linden S, Maksymowych WP, Mielants H, Olivieri I, Rudwaleit M. New criteria for inflammatory back pain in patients with chronic back pain: a real patient exercise by experts from the Assessment of SpondyloArthritis international Society (ASAS). Ann Rheum Dis. 2009;68(6):784-8.

28. Rudwaleit M, Metter A, Listing J, Sieper J, Braun J. Inflammatory back pain in ankylosing spondylitis: a reassessment of the clinical history for application as classification and diagnostic criteria. Arthritis Rheum. 2006;54(2):569-78.

29. Baraliakos X, Ghadir A, Fruth M, Kiltz U, Redeker I, Braun J. Which Magnetic Resonance Imaging Lesions in the Sacroiliac Joints Are Most Relevant for Diagnosing Axial Spondyloarthritis? A Prospective Study Comparing Rheumatologists' Evaluations With Radiologists' Findings. Arthritis Rheumatol. 2021;73(5):800-5.

30. Linden SVD, Valkenburg HA, Cats A. Evaluation of Diagnostic Criteria for Ankylosing Spondylitis. Arthritis & Rheumatism. 1984;27(4):361-8.

31. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, Braun J, Chou CT, Collantes-Estevez E, Dougados M, Huang F, Gu J, Khan MA, Kirazli Y, Maksymowych WP, Mielants H, Sorensen IJ, Ozgocmen S, Roussou E, Valle-Onate R, Weber U, Wei J, Sieper J. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis. 2009;68(6):777-83.

32. Poddubnyy D. Classification vs diagnostic criteria: the challenge of diagnosing axial spondyloarthritis. Rheumatology (Oxford). 2020;59(Suppl4):iv6-iv17.

33. Poddubnyy D, van Tubergen A, Landewe R, Sieper J, van der Heijde D, Assessment of SpondyloArthritis international S. Development of an ASAS-endorsed recommendation for the early referral of patients with a suspicion of axial spondyloarthritis. Ann Rheum Dis. 2015;74(8):1483-7.

34. van der Heijde D, Ramiro S, Landewe R, Baraliakos X, Van den Bosch F, Sepriano A, Regel A, Ciurea A, Dagfinrud H, Dougados M, van Gaalen F, Geher P, van der Horst-Bruinsma I, Inman RD, Jongkees M, Kiltz U, Kvien TK, Machado PM, Marzo-Ortega H, Molto A, Navarro-Compan V, Ozgocmen S, Pimentel-Santos FM, Reveille J, Rudwaleit M, Sieper J, Sampaio-Barros P, Wiek D, Braun J. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis. 2017;76(6):978-91.

35. Poddubnyy D, Diekhoff T, Baraliakos X, Hermann KGA, Sieper J. Diagnostic evaluation of the sacroiliac joints for axial spondyloarthritis: should MRI replace radiography? Ann Rheum Dis. 2022;81(11):1486-90.

36. Protopopov M, Proft F, Wichuk S, Machado PM, Lambert RG, Weber U, Juhl Pedersen S, Ostergaard M, Sieper J, Rudwaleit M, Baraliakos X, Maksymowych WP, Poddubnyy D. Comparing MRI and conventional radiography for the detection of structural changes indicative of axial spondyloarthritis in the ASAS cohort. Rheumatology (Oxford). 2022.

37. de Koning A, de Bruin F, van den Berg R, Ramiro S, Baraliakos X, Braun J, van Gaalen FA, Reijnierse M, van der Heijde D. 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(2):293-9.

38. Mandl P, Navarro-Compán V, Terslev L, Aegerter P, van der Heijde D, Agostino MA, Baraliakos X, Pedersen SJ, Jurik AG, Naredo E, Schueller-Weidekamm C, Weber U, Wick MC, Bakker PAC, Filippucci E, Conaghan PG, Rudwaleit M, Schett G, Sieper J, Tarp S, Marzo-Ortega H, Østergaard M. EULAR recommendations for the use of imaging in the diagnosis and management of spondyloarthritis in clinical practice. Annals of the Rheumatic Diseases. 2015;74(7):1327.

39. Dihlmann W. Morphological Basis for Numerous Interpretations of Sacroiliac X-Ray of Ankylosing-Spondylitis. Fortschr Rontg Neuen. 1976;124(4):389-90.

40. Creemers MC, Franssen MJ, van't Hof MA, Gribnau FW, van de Putte LB, van Riel PL. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis. 2005;64(1):127-9.

41. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, Dougados M, Hermann KG, Landewe R, Maksymowych W, van der Heijde D. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. Ann Rheum Dis. 2009;68 Suppl 2:ii1-44.

42. Maksymowych WP, Lambert RG, Baraliakos X, Weber U, Machado PM, Pedersen SJ, de Hooge M, Sieper J, Wichuk S, Poddubnyy D, Rudwaleit M, van der Heijde D, Landewe R, Eshed I, Ostergaard M. Data-driven definitions for active and structural MRI lesions in the sacroiliac joint in spondyloarthritis and their predictive utility. Rheumatology. 2021;60(10):4778-89.

43. Diekhoff T, Hermann KG, Greese J, Schwenke C, Poddubnyy D, Hamm B, Sieper J. Comparison of MRI with radiography for detecting structural lesions of the sacroiliac joint using CT as standard of reference: results from the SIMACT study. Ann Rheum Dis. 2017;76(9):1502-8.

44. Baraliakos X, Fruth M, Braun J. Bildgebung der Sakroiliakalgelenke. Zeitschrift für Rheumatologie. 2017;76(10):876-88.

45. Maksymowych WP, Lambert RGW, Ostergaard M, Pedersen SJ, Machado PM, Weber U, Bennett AN, Braun J, Burgos-Vargas R, de Hooge M, Deodhar AA, Eshed I, Jurik AG, Hermann KGA, Landewe RBM, Marzo-Ortega H, Navarro-Compan V, Poddubnyy D, Reijnierse M, Rudwaleit M, Sieper J, Van den Bosch FE, Van der Heijde D, van der Horst-Bruinsma IE, Wichuk S, Baraliakos X. MRI lesions in the sacroiliac joints of patients with spondyloarthritis: an update of definitions and validation by the ASAS MRI working group. Annals of the Rheumatic Diseases. 2019;78(11):1550-8.

46. Diekhoff T, Lambert R, Hermann KG. MRI in axial spondyloarthritis: understanding an 'ASAS-positive MRI' and the ASAS classification criteria. Skeletal Radiol. 2022.

47. Rudwaleit M, Jurik AG, Hermann KG, Landewe R, van der Heijde D, Baraliakos X, Marzo-Ortega H, Ostergaard M, Braun J, Sieper J. 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(10):1520-7.

48. Kiil RM, Mistegaard CE, Loft AG, Zejden A, Hendricks O, Jurik AG. Differences in topographical location of sacroiliac joint MRI lesions in patients with early axial spondyloarthritis and mechanical back pain. Arthritis Res Ther. 2022;24(1):75.

49. Bennett AN, McGonagle D, O'Connor P, Hensor EM, Sivera F, Coates LC, Emery P, Marzo-Ortega H. Severity of baseline magnetic resonance imaging-evident sacroiliitis and HLA-B27 status in early inflammatory back pain predict radiographically evident ankylosing spondylitis at eight years. Arthritis Rheum. 2008;58(11):3413-8.

50. Rudwaleit M, Listing J, Brandt J, Braun J, Sieper J. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor alpha blockers in ankylosing spondylitis. Ann Rheum Dis. 2004;63(6):665-70.

51. van den Berg R, de Hooge M, Rudwaleit M, Sieper J, van Gaalen F, Reijnierse M, Landewe R, Huizinga T, van der Heijde D. ASAS modification of the Berlin algorithm for diagnosing axial spondyloarthritis: results from the SPondyloArthritis Caught Early (SPACE)-cohort and from the

Assessment of SpondyloArthritis international Society (ASAS)-cohort. Annals of the Rheumatic Diseases. 2013;72(10):1646-53.

52. Proft F, Spiller L, Redeker I, Protopopov M, Rodriguez VR, Muche B, Rademacher J, Weber AK, Luders S, Torgutalp M, Sieper J, Poddubnyy D. Comparison of an online self-referral tool with a physician-based referral strategy for early recognition of patients with a high probability of axial spa. Semin Arthritis Rheu. 2020;50(5):1015-21.

53. Redeker I, Callhoff J, Hoffmann F, Haibel H, Sieper J, Zink A, Poddubnyy D. Determinants of diagnostic delay in axial spondyloarthritis: an analysis based on linked claims and patient-reported survey data. Rheumatology (Oxford). 2019;58(9):1634-8.

54. Feldtkeller E. [Age at disease onset and delayed diagnosis of spondyloarthropathies]. Z Rheumatol. 1999;58(1):21-30.

55. Poddubnyy D. Axiale Spondyloarthritis (Morbus Bechterew) – Update 2021. Dtsch Med Wochenschr. 2021;146(03):185-90.

56. Machado P, Landewe R, Lie E, Kvien TK, Braun J, Baker D, van der Heijde D, Assessment of SpondyloArthritis international S. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis. 2011;70(1):47-53.

57. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol. 1994;21(12):2286-91.

58. Castillo C, Steffens T, Sim L, Caffery L. The effect of clinical information on radiology reporting: A systematic review. J Med Radiat Sci. 2021;68(1):60-74.

59. Loy CT, Irwig L. Accuracy of diagnostic tests read with and without clinical information: a systematic review. JAMA. 2004;292(13):1602-9.

60. Song KS, Song HH, Park SH, Ahn KJ, Yang IK, Byun JY, Jeon JS, Kim JY, Kim BS, Lim GY. Impact of clinical history on film interpretation. Yonsei Med J. 1992;33(2):168-72.

61. Brandt HC, Spiller I, Song IH, Vahldiek JL, Rudwaleit M, Sieper J. Performance of referral recommendations in patients with chronic back pain and suspected axial spondyloarthritis. Ann Rheum Dis. 2007;66(11):1479-84.

62. Sudol-Szopinska I, Jurik AG, Eshed I, Lennart J, Grainger A, Ostergaard M, Klauser A, Cotten A, Wick MC, Maas M, Miese F, Egund N, Boutry N, Rupreht M, Reijnierse M, Oei EH, Meier R, O'Connor P, Feydy A, Mascarenhas V, Plagou A, Simoni P, Platzgummer H, Rennie WJ, Mester A, Teh J, Robinson P, Guglielmi G, Astrom G, Schueller-Weiderkamm C. Recommendations of the ESSR Arthritis Subcommittee for the Use of Magnetic Resonance Imaging in Musculoskeletal Rheumatic Diseases. Semin Musculoskelet Radiol. 2015;19(4):396-411.

63. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377-81.

64. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, Duda SN, Consortium RE. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019;95:103208.

65. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.

66. Kiltz U, Baraliakos X, Karakostas P, Igelmann M, Kalthoff L, Klink C, Krause D, Schmitz-Bortz E, Florecke M, Bollow M, Braun J. Do patients with non-radiographic axial spondylarthritis differ from patients with ankylosing spondylitis? Arthritis Care Res (Hoboken). 2012;64(9):1415-22.

67. Joish UK, Sahni H, Reddy RH, Sreekanth V. Clinical radiology: Going an extra mile for patient care. Indian J Radiol Imaging. 2018;28(2):263-6.

68. Schon F, Sinzig R, Walther F, Radosa CG, Nebelung H, Eberlein-Gonska M, Hoffmann RT, Kuhn JP, Blum SFU. Value of Clinical Information on Radiology Reports in Oncological Imaging. Diagnostics (Basel). 2022;12(7).

69. Hattori S, Yokota H, Takada T, Horikoshi T, Takishima H, Mikami W, Mukai H, Funakoshi H, Naganuma T, Yano T, Terada K, Inokuchi R, Tsuchiya S, Akutsu A, Saiga A, Kubota Y, Wada T, Takada A, Fujimoto H, Uno T. Impact of clinical information on CT diagnosis by radiologist and subsequent clinical management by physician in acute abdominal pain. Eur Radiol. 2021;31(8):5454-63.

70. Chilanga CC, Olerud HM, Lysdahl KB. The value of referral information and assessment - a cross sectional study of radiographers' perceptions. BMC Health Serv Res. 2022;22(1):893.

71. Griscom NT. A suggestion: look at the images first, before you read the history. Radiology. 2002;223(1):9-10.

72. Fatahi N, Krupic F, Hellstrom M. Quality of radiologists' communication with other clinicians--As experienced by radiologists. Patient Educ Couns. 2015;98(6):722-7.

73. Poddubnyy D, Weineck H, Diekhoff T, Redeker I, Gobejishvili N, Llop M, Rodriguez VR, Proft F, Protopopov M, Haibel H, Sieper J, Hermann KGA. Clinical and imaging characteristics of osteitis condensans ilii as compared with axial spondyloarthritis. Rheumatology. 2020;59(12):3798-806.

74. van der Heijde D, Braun J, Deodhar A, Baraliakos X, Landewe R, Richards HB, Porter B, Readie A. Modified stoke ankylosing spondylitis spinal score as an outcome measure to assess the impact of treatment on structural progression in ankylosing spondylitis. Rheumatology (Oxford). 2019;58(3):388-400.

75. Diekhoff T, Greese J, Sieper J, Poddubnyy D, Hamm B, Hermann KGA. Improved detection of erosions in the sacroiliac joints on MRI with volumetric interpolated breath-hold examination (VIBE): results from the SIMACT study. Annals of the Rheumatic Diseases. 2018;77(11):1585-9.

76. Diekhoff T, Eshed I, Radny F, Ziegeler K, Proft F, Greese J, Deppe D, Biesen R, Hermann KG, Poddubnyy D. Choose wisely: imaging for diagnosis of axial spondyloarthritis. Ann Rheum Dis. 2022;81(2):237-42.

Appendix

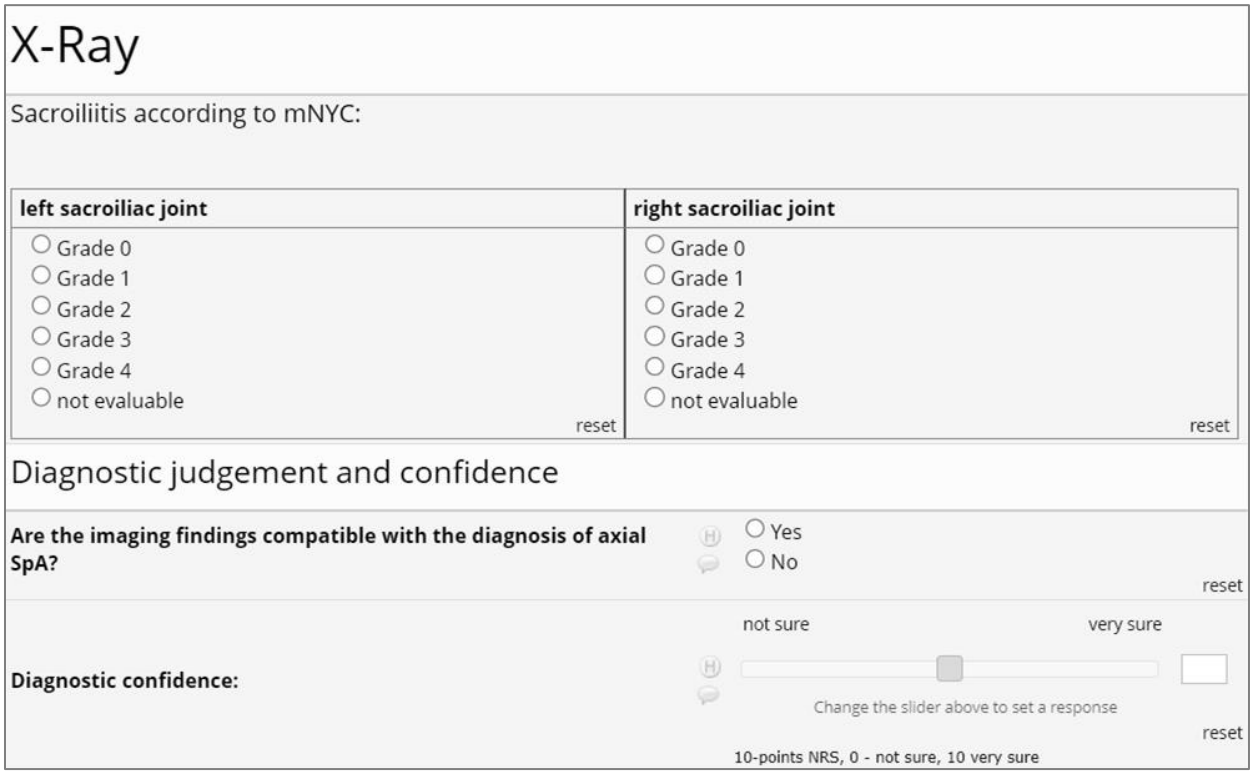


Abbildung A 1 Befundmaske der Runde 1 (Röntgen ohne klinische Informationen) zur Auswertung der Bilder durch die Radiolog*innen

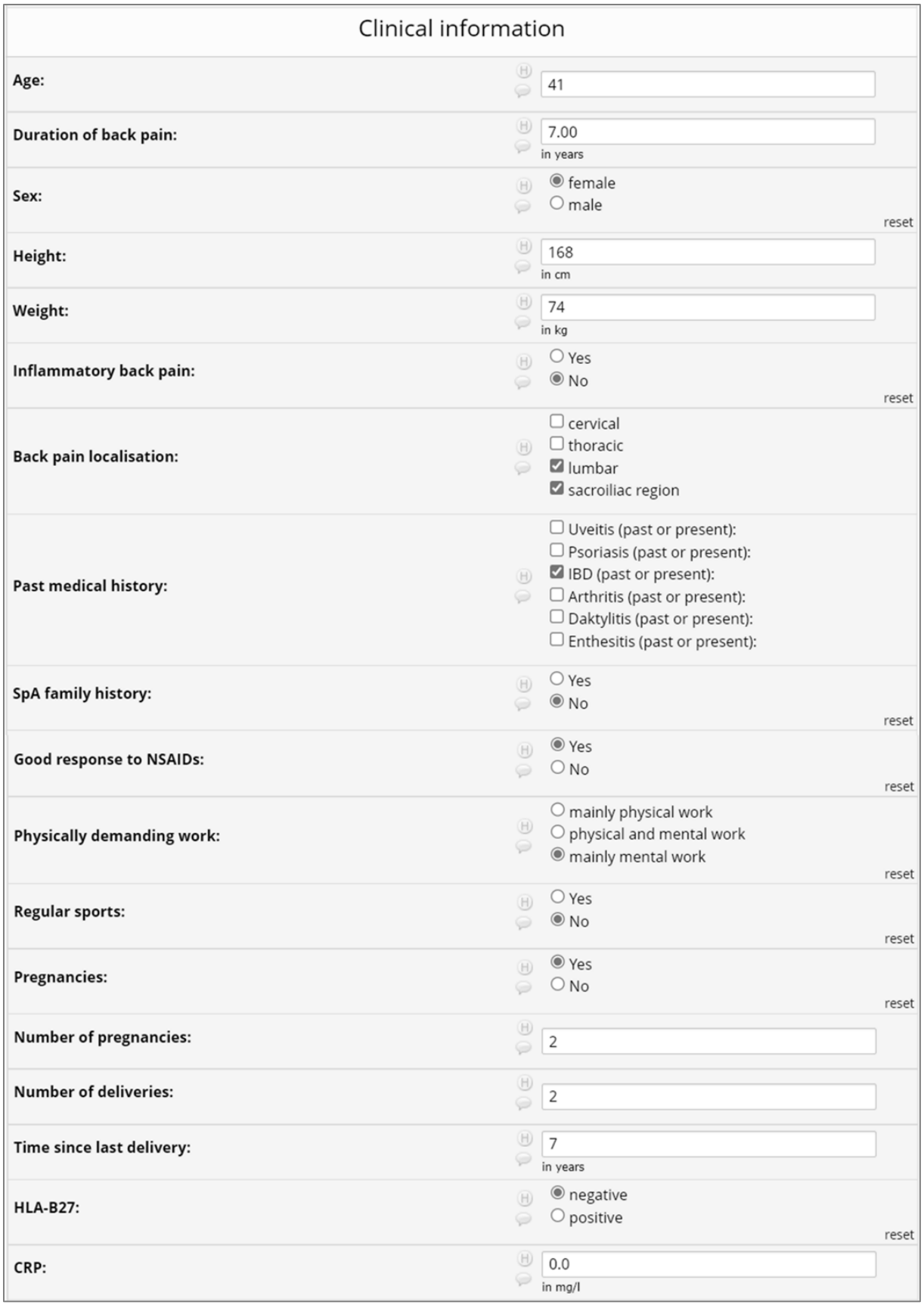


Abbildung A 2 Liste der klinischen Informationen (fiktives Beispiel), die den Radiolog*innen zur Verfügung gestellt wurde.

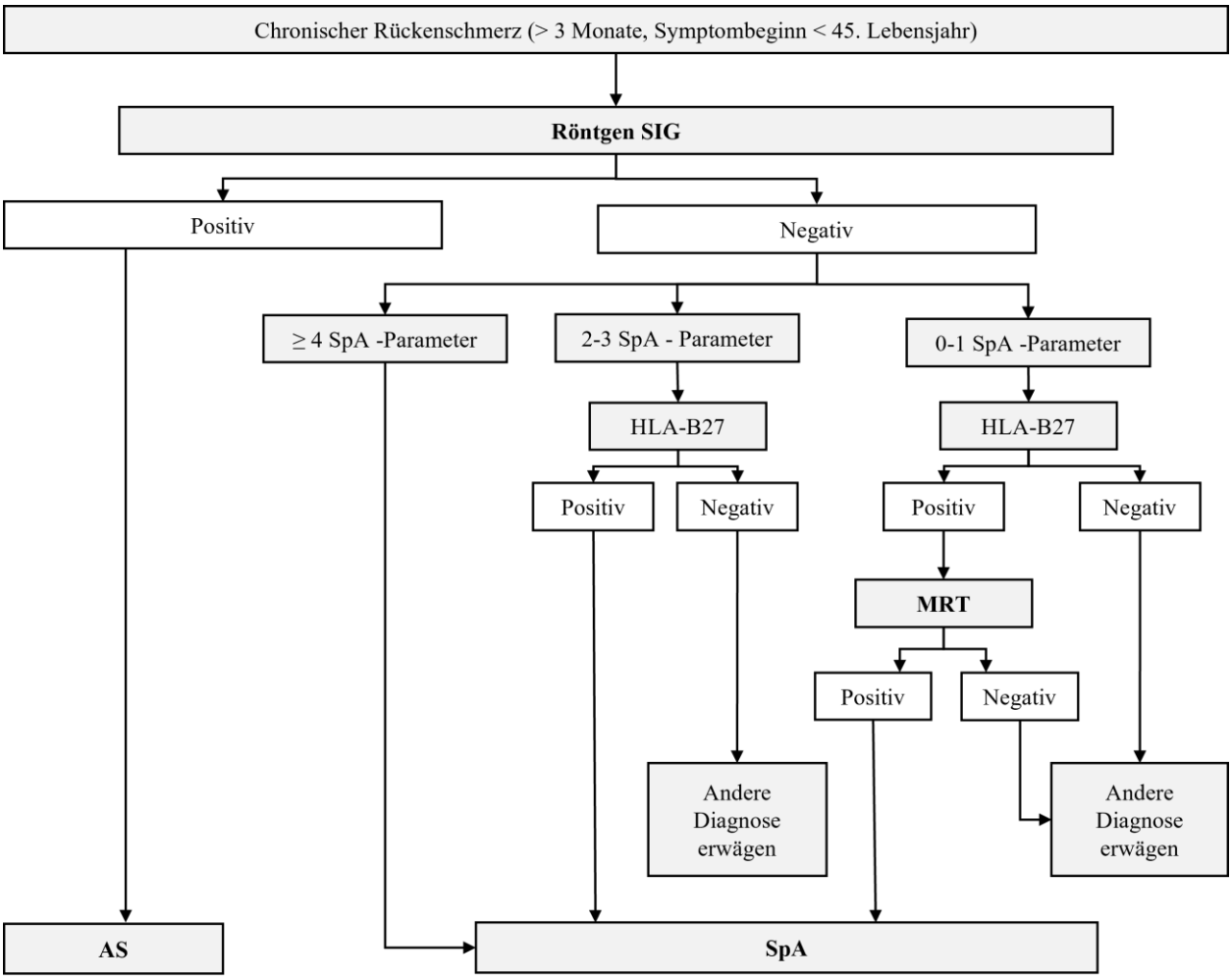


Abbildung A 3: **ASAS modifizierter Berlin Algorithmus** (übersetzt und modifiziert nach: van den Berg et al. (51))

Imaging request	
1	The referring physician should communicate important clinical information when requesting imaging exams. This clinical information should include the patient's age, sex and HLA-B27 status.
2	Requests for imaging should include current or past history of back pain, its duration, localization, and inflammatory features, whether present or not. For follow-up exams, a change in clinical symptoms should be indicated.
3	Radiologists should be informed if the patient undertakes physically demanding activities or has history of childbirth (number of children and date of most recent childbirth).
4	Radiologists should have access to previous exam images for comparison or to the respective reports if those are not available.
5	The referral should include possible contraindications to certain types of imaging or contrast medium.
6	The referring physician should indicate the suspected clinical diagnosis and possible alternative explanations for the symptoms, whether SpA was previously diagnosed, and if the exam is requested for primary diagnosis, to assess disease activity or treatment response.

Imaging report	
Clinical data	
1	The report should start by summarising essential clinical information, including the patient's age, sex, a summary of symptoms, the suspected diagnosis, whether the exam was requested for primary diagnosis or follow-up, and what imaging was available for comparison.
Technical data	
2a	Radiography: The report should include the number of images, types of projections, and the patient's positioning.
2b	MRI: The report should include the applied field strength and sequences with slice orientation and thickness, if fat suppression was applied, and whether and what type of contrast medium was administered.
2c	CT: The report should include the patient's position, reconstructions' orientation and slice thickness, and a general indicator for the radiation dose (e.g., dose length product).
3	The anatomical coverage of the exam should be indicated.
4	The report should include a general statement about image quality and complications from imaging, particularly if the exam or its interpretation is affected.
Report	
5a	SIJ: Bone marrow oedema/osteitis, erosions and fat lesions are significant findings that the report should list semi-quantitatively with their localization specified. Their absence should be stated clearly.
5b	SIJ: The report should include if other active or structural lesions are present. Structural lesions should be reported per individual bone. The radiologists can summarize the absence of those active or structural lesions in the report.
6a	Spine: The report should semi-quantitatively list bone marrow oedema/osteitis at vertebral corners. All other active and structural lesions should be mentioned if present.
6b	Spine: The location of the findings mentioned above is essential for clinical correlation, and it should be stated at the level of the individual vertebra or discovertebral unit.
7	Findings unrelated to SpA but of potential clinical importance should be mentioned when present. These include for example gas inside the joint ("vacuum phenomenon"), osteophytes, transitional vertebrae, anatomical variations, and spinal malposition.
Conclusion	
8	The radiologist should state clearly if findings are compatible with SpA, based on the images and clinical information available. The conclusion should provide whether there is active inflammation or structural changes with the most prominent lesions, and give an indication of the confidence in interpretation of the findings.
9	Based on the exam findings, differential diagnoses and their probability should be mentioned, especially if more likely than SpA.
10	If the exam findings are inconclusive, radiologists should suggest further imaging according to their expertise.
11	If the exam is indicative of SpA and a rheumatologist did not request the imaging investigation, the report should recommend referral to a rheumatologist for further assessment.

Abbildung A 4 ASAS-Empfehlungen: Radiologischer Befund: Empfehlungen für die inhaltliche Ausgestaltung von a) der Anforderung an die Rheumatolog*innen bei V.a. axSpA und b) der Befundbeschreibung. (aus Diekhof et. al (1))

Eidesstattliche Versicherung

Ich, Tim Pohlner, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema: „Der Einfluss der klinischen Information auf die radiologische Befundung bei Patient*innen mit Verdacht auf axiale Spondyloarthritis“ (engl. „Axial spondyloarthritis: The role of clinical information in imaging evaluation“) selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren/innen beruhen, sind als solche in korrekter Zitierung kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) werden von mir verantwortet.

Ich versichere ferner, dass ich die in Zusammenarbeit mit anderen Personen generierten Daten, Datenauswertungen und Schlussfolgerungen korrekt gekennzeichnet und meinen eigenen Beitrag sowie die Beiträge anderer Personen korrekt kenntlich gemacht habe (siehe Anteilserklärung). Texte oder Textteile, die gemeinsam mit anderen erstellt oder verwendet wurden, habe ich korrekt kenntlich gemacht.

Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Erstbetreuer/in, angegeben sind. Für sämtliche im Rahmen der Dissertation entstandenen Publikationen wurden die Richtlinien des ICMJE (International Committee of Medical Journal Editors; www.icmje.og) zur Autorenschaft eingehalten. Ich erkläre ferner, dass ich mich zur Einhaltung der Satzung der Charité – Universitätsmedizin Berlin zur Sicherung Guter Wissenschaftlicher Praxis verpflichte.

Weiterhin versichere ich, dass ich diese Dissertation weder in gleicher noch in ähnlicher Form bereits an einer anderen Fakultät eingereicht habe.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§§156, 161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

_____ _____

Datum Unterschrift

Anteilserklärung der erfolgten Publikation

Ich hatte folgenden Anteil an den folgenden Publikationen:

Pohlner T, Deppe D, Ziegeler K, Proft F, Protopopov M, Rademacher J, Rios Rodriguez V, Torgutalp M, Braun J, Diekhoff T, Poddubnyy D. Diagnostic accuracy in axial spondyloarthritis: a systematic evaluation of the role of clinical information in the interpretation of sacroiliac joint imaging. RMD Open. 2024;10(2).

Als Erstautor dieser Publikation habe ich die folgenden Anteile:

- Planung des Studiendesigns gemeinsam mit Professor Puddubnyy
- Vorauswahl des Patient*innenkollektivs (Prüfung auf Vollständigkeit der Daten)
- Erstellung der Befundmasken zur Auswertung durch die Radiolog*innen (elektronisch über die Online-Plattform RedCAP)
- Statistische Auswertung der erhobenen Befunde
- Interpretation der Ergebnisse gemeinsam mit den anderen Autor*innen
- Anfertigung der Tabellen 1 bis 3 sowie der Abbildung 1. Anfertigung der Tabelle 4 gemeinsam mit den anderen Autor*innen
- Gemeinsame Anfertigung des Manuskripts mit den anderen Autor*innen

_____ _____

Datum Unterschrift

Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Publikationsliste

Publikation 1 (von 1):

Pohlner T, Deppe D, Ziegeler K, Proft F, Protopopov M, Rademacher J, Rios Rodriguez V, Torgutalp M, Braun J, Diekhoff T, Poddubnyy D. Diagnostic accuracy in axial spondyloarthritis: a systematic evaluation of the role of clinical information in the interpretation of sacroiliac joint imaging. RMD Open. 2024;10(2).

Danksagung

An dieser Stelle möchte ich allen beteiligten Personen meinen großen Dank aussprechen, die mich bei der Anfertigung und Bearbeitung meiner Doktorarbeit unterstützt haben.

Mein besonderer Dank gilt Professor Poddubnyy für die hervorragende Betreuung bei der Umsetzung der gesamten Arbeit sowie für die verlässliche Hilfestellung bei jeglichen Fragen oder Problemen.

Außerdem möchte ich mich bei Dr. Katharina Ziegeler, Dr. Torsten Diekhoff und Dr. Dominik Deppe für die zuverlässige und sorgfältige Auswertung der Röntgen- und MRT-Bilder bedanken.

Darüber hinaus möchte ich dem gesamten Team der Rheumatologie der Charité am Campus Benjamin Franklin meinen Dank für die Unterstützung bei verschiedenen Herausforderungen dieser Arbeit bedanken, besonders bei Dr. Judith Rademacher für die Einführung in das Programm RedCAP.

Zuletzt gilt mein großer Dank all meinen Freunden und meiner Familie, insbesondere Ena und Ronja für die großartige Hilfe bei statistischen Fragen, meinen Eltern, meinen Schwiegereltern und natürlich Caro für die riesige Unterstützung und Motivation bei dieser Arbeit.