Imaging outcomes for axial spondyloarthritis and sensitivity to change: a five-year analysis of the DESIR cohort

Sepriano, A.; Ramiro, S.; Heijde, D. van der; Dougados, M.; Claudepierre, P.; Feydy, A.; ... ; Landewe, R.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 251–258
DOI 10.1002/acr.24459

American College of Rheumatology
Empowering Rheumatology Professionals

Imaging Outcomes for Axial Spondyloarthritis and Sensitivity to Change: A Five-Year Analysis of the DESIR Cohort

Alexandre Sepriano,[1] Sofia Ramiro,[2] Désirée van der Heijde,[3] Maxime Dougados,[4] Pascal Claudepierre,[5] Antoine Feydy,[6] Monique Reijnierse,[3] Damien Loeuille,[7] and Robert Landewé[8]

Objective. To compare the sensitivity to change of different imaging scoring methods in patients with early axial spondyloarthritis (SpA).

Methods. Patients from the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort fulfilling the Assessment of SpondyloArthritis international Society criteria for axial SpA were included. Radiographs and magnetic resonance imaging (MRI) of the sacroiliac (SI) joints and spine were obtained at baseline, 1, 2, and 5 years. Each image was scored by 2 or 3 readers in 3 separate reading waves. The rate of change of outcomes measuring inflammation of the spine and SI joints (e.g., Spondyloarthritis Research Consortium of Canada [SPARCC] score) and structural damage on MRI (e.g., ≥3 fatty lesions) and radiographs (e.g., modified New York grading) was assessed using multilevel generalized estimating equation models (taking all readers and waves into account). To allow comparisons across outcomes, rates were standardized (difference between the individual's value and the population mean divided by the SD).

Results. In total, 345 patients were included. Inflammation detected on MRI of the SI joints (MRI-SI joints) (standardized rate range –0.278, –0.441) was more sensitive to change compared to spinal inflammation (range –0.030, –0.055). Structural damage in the SI joints showed a higher standardized rate of change on MRI-SI joints (range 0.015, 0.274) compared to radiography of the SI joints (range 0.043, 0.126). MRI-SI joints damage defined by ≥3 fatty lesions showed the highest sensitivity to change (0.274). Spinal structural damage slowly progressed over time with no meaningful difference between radiographic (range 0.037, 0.043) and MRI structural outcomes (range 0.008, 0.027).

Conclusion. Structural damage assessed in pelvic radiographs has low sensitivity to change, while fatty lesions detected on MRI-SI joints are a promising alternative. In contrast, MRI of the spine is not better than radiography of the spine in detecting structural changes in patients with early axial SpA.

INTRODUCTION

Several imaging outcomes have been developed to assess inflammation and structural damage over time in patients with axial spondyloarthritis (SpA). A recent systematic literature review informing the European Alliance of Associations for Rheumatology recommendations for the use of imaging in the diagnosis and management of SpA in clinical practice identified several studies testing

The DESIR study is supported by an unrestricted grant from Pfizer and is conducted as a program of the Hospitalier de Recherche Clinique with Assistance Publique Hopitaux de Paris as the sponsor and under the umbrella of the French Society of Rheumatology, which supports the cohort. The DESIR cohort is conducted under the Assistance Publique Hopitaux de Paris via the Clinical Research Unit Paris Centre and under the umbrella of the French Society of Rheumatology and Institut national de la sante et de la recherche medicale (INSERM). Database management is performed within the Department of Epidemiology and Biostatistics. Dr. Sepriano's work was supported by the Fundação para a Ciência e Tecnologia (doctoral grant SFRH/BD/108246/2015).

[1]Alexandre Sepriano, MD: Leiden University Medical Center, Leiden, The Netherlands, and NOVA Medical School, Universidade Nova de Lisboa, Lisbon, Portugal; [2]Sofia Ramiro, MD, PhD: Leiden University Medical Center, Leiden, and Zuyderland Medical Center, Heerlen, The Netherlands; [3]Désirée van der Heijde, MD, PhD, Monique Reijnierse, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; [4]Maxime Dougados, MD: Department of Rheumatology, Hôpital Cochin, Assistance Publique, Hôpitaux de Paris, INSERM (U1153), and Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-Cité, Paris, France; [5]Pascal Claudepierre, MD: Department of Rheumatology, Hôpital Henri-Mondor, Assistance Publique-Hôpitaux de Paris, and Université Paris Est Créteil, Créteil, France; [6]Antoine Feydy, MD, PhD: Radiology B Department, Cochin Hôpital Paris and Descartes University, Paris, France; [7]Damien Loeuille, MD, PhD: Department of Rheumatology, Hôpital Brabois, Nancy, France; [8]Robert Landewé, MD, PhD: Amsterdam University Medical Center, Amsterdam, and Zuyderland Medical Center, Heerlen, The Netherlands.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Alexandre Sepriano, MD, Leiden University Medical Center, Department of Rheumatology, Albinusdreef 2, PO Box 9600, 2300 RC Leiden, The Netherlands. Email: alexsepriano@gmail.com.

Submitted for publication March 9, 2020; accepted in revised form September 15, 2020.

SIGNIFICANCE & INNOVATIONS

- Several imaging outcomes are available to measure inflammation and damage over time in patients with axial spondyloarthritis (SpA); however, direct comparisons of their sensitivity to change are scarce, especially in early disease.
- In early axial SpA, outcomes of inflammation measured on magnetic resonance imaging (MRI) are more sensitive to change on the sacroiliac (SI) joints than on the spine.
- MRI of the SI joints is more sensitive in capturing change in structural damage, especially fatty lesions, than pelvic radiographs, while MRI of the spine is not better than spinal radiographs in detecting structural changes in patients with early axial SpA.
- Results from this study may help in prioritizing imaging scoring methods in subsequent observational or interventional studies in early axial SpA.

the utility of magnetic resonance imaging (MRI) and radiographs of the sacroiliac (SI) joints and spine on monitoring disease activity and structural damage over time (1). However, these studies mostly assessed only 1 score each and focused on comparing imaging to clinical measures of disease activity, disability, and mobility, which means that they mostly addressed their validity.

In addition to validity, in order to prioritize imaging outcomes measuring similar aspects of the disease (i.e., inflammation or structural damage), the other aspects of the Outcome Measures in Rheumatology (OMERACT) filter, namely discrimination (sensitivity to change and reliability) and feasibility, should also be taken into account (2). However, direct comparisons of the discriminative ability and feasibility of imaging outcomes in axial SpA have been seldom performed, and almost only in later phases of the disease (radiographic axial SpA) (3–5). An exception to this is the comparison of the different spinal radiographic scoring methods performed in the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort and previously reported by our team (6).

A better understanding of which imaging findings (reflecting inflammation or structural damage), imaging modality (MRI or radiographs), and anatomic location (SI joints or spine) are most informative to monitor axial changes in the entire spectrum of axial SpA (also including nonradiographic axial SpA) over time is still a major unmet need. We aimed to compare the sensitivity to change of different MRI and radiographic scoring methods in patients with early axial SpA.

PATIENTS AND METHODS

Patients and study design. Five-year data from patients with early axial SpA from the DESIR cohort have been used (ClinicalTrials.gov identifier: NCT01648907) (7). Patients had to fulfill the Assessment of SpondyloArthritis international Society (ASAS) criteria for axial SpA and to have ≥1 radiograph and/or MRI reading available during the 5-year follow-up period to be included in the current study. The database used for the current analysis was locked on June 20, 2016. The study was approved by the appropriate local medical ethical committees. All patients provided signed informed consent upon participation.

Imaging scoring procedures. Radiographs of the SI joints and spine and MRIs of the SI joints (MRI-SI joints) and spine (MRI-spine) were obtained at baseline, 1, 2, and 5 years. Each image was independently scored in 3 reading waves by trained central readers blinded to chronology, clinical data, and to the results of other imaging modalities. In wave 1, baseline images were scored by 2 readers and 1 adjudicator (in case of disagreement). In wave 2, images from baseline, 1, and 2 years were also scored by 2 readers and 1 adjudicator. In wave 3, images from baseline, 2, and 5 years were scored by 3 central readers. Readers and adjudicators varied across modalities and waves (8) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24459/abstract). By protocol, radiographs were performed in all 25 participating centers at each time point, but MRIs were only performed in all centers at baseline, while MRIs at 1, 2, and 5 years were only obtained in 9 centers from Paris.

Inflammation outcomes. Inflammation on MRI-SI joints was assessed using the ASAS definition (positive/negative) and the Spondyloarthritis Research Consortium of Canada (SPARCC) score (range 0–72) (9–11). Bone marrow edema (BME) on MRI-spine was defined according to the ASAS definition (≥3 vertebral corner lesions; yes/no) (12). In addition, a cutoff of 5 vertebral corner BME lesions (typical of axial SpA and present in ≥2 consecutive slices) was also assessed according to the Canada–Denmark method, as it has been shown to be highly specific of axial SpA (13). The total spine SPARCC score (range 0–414) and Berlin score (range 0–69) were used as continuous inflammatory outcomes (3,14).

Structural outcomes. Structural damage on radiography of the SI joints was assessed according to the modified New York (mNY) system as continuous (range 0–8) and as a binary (positive/negative) score (15). Two additional binary definitions were assessed: worsening of ≥1 grade in ≥1 SI joints (yes/no); and worsening of ≥1 grade in ≥1 SI joints, with a 5-year grade ≥2 in the worsened joint (yes/no) (16).

An adaptation of the MRI-SI joints structural score by Weber et al, previously described by our team (17), was used to define individual structural lesions on MRI-SI joints (18). In summary, fatty lesions, erosions, and ankylosis/partial ankylosis are scored as originally described. Sclerosis was added. Fatty lesions, erosions, and sclerosis were marked as present if seen on ≥2

consecutive slices (maximum 5 lesions in 6 slices per each of the 8 quadrants in both SI joints). Ankylosis or partial ankylosis was considered present if seen on a single slice. Partial ankylosis and ankylosis cannot occur simultaneously in a quadrant, and ankylosis always involves 2 quadrants; therefore, the corresponding scoring range is 0–24. In the absence of a formal definition of presence of structural damage on MRI-SI joints, we considered 3 definitions previously shown most discriminatory in early axial SpA: ≥5 fatty lesions and/or erosions; ≥3 erosions; and ≥3 fatty lesions (13). Continuous structural lesions on MRI-SI joints were defined as number of fatty lesions and/or erosions (range 0–80), number of erosions (range 0–40), number of fatty lesions (range 0–40), and total number of lesions with (range 0–144) and without (range 0–104) sclerosis.

Structural lesions on radiography of the spine were assessed as the presence of ≥1 syndesmophyte (yes/no) and by using the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS; range 0–72) (19). Structural lesions on MRI-spine were scored according to the Canada–Denmark method (20,21). In the absence of a formal definition, we defined structural damage as ≥5 fatty lesions, also previously shown to be highly specific for axial SpA (13). The total number of structural lesions (fatty lesions, erosions, bone spurs, ankylosis) (range 0–322) was assessed, as well as the total number of fatty lesions, erosions, and bone spurs (range 0–92 for all).

A detailed description of all scores is provided in Supplementary Tables 2–10, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24459/abstract. The interreader reliability of the radiographic and MRI outcomes used in this study has been reported in detail elsewhere (6,17) and is summarized in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24459/abstract.

Statistical analysis. The baseline value for each outcome was defined by a combination algorithm of the scores from the 3 readers from wave 3 (agreement between ≥2 of 3 for binary, and a mean of 3 readers for continuous outcomes). The rate of change of each outcome was analyzed using generalized estimating equations (GEEs), with time in years as the explanatory variable of interest. Each outcome was analyzed per patient, per time point and per individual reader, and the yearly rate of change estimated using so-called integrated analysis, including all patients with ≥1 score from ≥1 reader from ≥1 reading wave. Different to traditional measures of sensitivity to change (e.g., Cohen's effect size), this method, which we have previously explained in detail (8), appropriately handles the multilevel data structure of our data. All patients had to have ≥1 score from all outcomes, thus ensuring that the same patients are used across all analyses. All variables were standardized. A standardized variable (metric free) was defined at the patient level as the difference between the individual's value and the population mean divided by the population SD. Each standardized variable has a mean of 0 and a variance of 1 and reads as the number of SD above (positive) or below (negative) the mean.

In addition, the relative standardized rate of change (i.e., the standardized yearly rate of change of an outcome divided by the corresponding rate of a reference imaging outcome) was calculated. For this calculation, a value >1 means larger sensitivity, and a value <1 lower sensitivity compared to the reference (the further away from 1, the larger the difference). Three types of references were defined: 1) inflammation common reference (comparing all inflammation outcomes to sacroiliitis on MRI-SI joints [ASAS definition]); 2) structural common reference (comparing all structural outcomes to sacroiliitis on radiography of the SI joints [mNY]); and 3) modality reference (comparing outcomes to a reference within each modality and anatomic site).

Goodness-of-fit statistics (quasi-likelihood under the independence model criterion [QIC]) were used to get an impression on how much of the outcome variability was explained by each model. Different transformations of time were tested to assess which one yielded the lowest QIC (better fit). A nonlinear model was chosen if it best fit the data and if the nonlinear factor (e.g., quadratic term) added to the model was significant ($P < 0.05$). Stata, version 15.1, was used for the analyses.

RESULTS

Baseline characteristics. In total, 345 patients were included (mean ± SD symptom duration 1.6 ± 0.9 years; 53% were male patients, and 89% HLA–B27 positive [Table 1]). Baseline inflammation on MRI was more frequently present at the SI joints (active sacroiliitis: 39%) than at the spine level (BME ≥5 lesions: 6%) (Table 2). Structural damage at baseline was limited in the SI joints (21% mNY positive) and even more in the spine (≥1 syndesmophyte: 6%) (Table 3).

Sensitivity to change of the different imaging outcomes. Inflammation on MRI-SI joints showed a higher sensitivity to change than on MRI-spine, the latter remaining essentially unchanged over time. This was true for the dichotomous ASAS MRI-SI joints score (standardized yearly rate of change –0.278) and especially for the continuous SPARCC score (standardized yearly rate of change –0.441), while the standardized yearly rates of change for MRI-spine ranged only between –0.030 and –0.055 (Table 2). The differences between SI joints and spine inflammation outcomes become especially evident with the relative standardized rate of change. Compared to the ASAS definition of a positive MRI-SI joints (inflammation common reference, i.e., a value of 1), all inflammation outcomes in the spine were much less sensitive to change (range of relative standardized rates 0.094–0.531; i.e., all values far below 1).

Structural damage in the SI joints increased over time but with a larger yearly rate on MRI-SI joints (standardized rate range 0.015–0.274) compared to radiography of the SI joints

Table 1. Patient and disease characteristics at baseline and during follow-up*

Characteristic	Baseline (n = 345)	1 year (n = 345)	2 years (n = 342)	5 years (n = 320)
Age at baseline, mean ± SD years	31.0 ± 7.0	–	–	–
Male sex	183 (53)	–	–	–
Symptoms duration, mean ± SD years	1.6 ± 0.9	–	–	–
Current smokers†	135 (39)	127 (39)	118 (37)	92 (34)
HLA–B27	307 (89)	–	–	–
Radiographic sacroiliitis (mNY)‡	73 (21)	NA	68 (23)	68 (27)
BASDAI score, mean ± SD (range 0–10)†	4.1 ± 2.0	3.2 ± 2.2	3.1 ± 2.2	2.9 ± 2.0
ASDAS-CRP score, mean ± SD‡	2.6 ± 1.0	2.1 ± 0.9	2.0 ± 0.9	2.0 ± 0.9
Elevated CRP (≥6 mg/liter)‡	109 (33)	64 (20)	69 (22)	57 (22)
BASFI score, mean ± SD (0–10)†	2.7 ± 2.2	2.1 ± 2.1	2.1 ± 2.2	2.0 ± 2.0
TNFi treatment‡	0 (0)	76 (24)	94 (29)	111 (42)
NSAID treatment†	329 (95)	250 (77)	216 (68)	180 (66)

* Values are the number (%) unless indicated otherwise. ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; CRP = C-reactive protein; mNY = modified New York criteria (scored in wave 3); NA = not applicable (imaging in wave 3 is only scored at baseline, 2 years, and 5 years); NSAID = nonsteroidal anti-inflammatory drugs; TNFi = tumor necrosis factor inhibitors.
† Missing data <15% in each visit.
‡ Missing data <20% in each visit.

(standardized rate range 0.043–0.126) (Table 3). Three or more fatty lesions on MRI-SI joints was the SI joints structural outcome with highest sensitivity to change (standardized rate 0.274; relative rate of 6.227 comparing to mNY). On the contrary, ≥3 erosions on MRI-SI joints was the least sensitive (standardized rate 0.015) of all SI joints structural outcomes (including both MRI-SI joints and radiography of the SI joints). Importantly, ≥3 fatty lesions alone was slightly more sensitive to change than combining fatty lesions with erosions, i.e., ≥5 fatty lesion and/or erosions (relative rate of 1.151 for the former compared to the latter).

Among the radiography of the SI joints structural outcomes, worsening of ≥1 grade in ≥1 SI joints and worsening of ≥1 grade in ≥1 SI joints, with a 5-year grade ≥2 in the worsened joint, were far more sensitive to change compared to the mNY binary definition as the modality reference (relative rate 2.864 and 2.705, respectively). Of note, the mNY continuous grading and the mNY binary score had comparable sensitivity to change (relative rate of the continuous versus the reference binary score = 0.977).

Overall, the standardized yearly rate of change of the spinal radiographic outcomes (range 0.037–0.043) was higher as compared to MRI-spine structural outcomes (range 0.012–0.027) (Table 3), although all are relatively low. Among MRI-spine outcomes, the total number of bone spurs was the outcome that most captured change (standardized rate 0.027; and relative rate of 2.077 compared to ≥5 fatty lesions, i.e., the modality reference). Yet, the best MRI-spine outcome is still less sensitive to change as compared to radiography of the spine outcomes, with

Table 2. Baseline score and standardized yearly rate of change (ROC) of inflammatory imaging outcomes over 5 years of follow-up in patients with early axial spondyloarthritis (SpA) fulfilling the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axial SpA*

Imaging outcomes	Baseline score (range 334–344)†	Standardized ROC per year‡	Relative standardized ROC§	Relative standardized ROC per modality and anatomic site
Inflammatory lesions (MRI of the SI joints)¶				
Sacroiliitis (ASAS criteria), no. (%)	134 (39.2)	–0.278#	1	1
SPARCC SI joint score (range 0–72)	4.7 ± 7.9	–0.441#	1.586	1.586
Inflammatory lesions (MRI of the spine)**				
BME ≥3 lesions, no. (%)	32 (9.4)	–0.032	0.319	1
BME ≥5 lesions, no. (%)	19 (5.6)	–0.030	0.094	0.938
23-DVU SPARCC spine score (range 0–414)	2.6 ± 7.7	–0.050	0.531	1.563
Berlin spine score (range 0–69)	0.9 ± 2.7	–0.055	0.104	1.719

* Values are the mean ± SD unless indicated otherwise. BME = bone marrow edema; DVU = discovertebral unit; MRI = magnetic resonance imaging; SI = sacroiliac; SPARCC = Spondyloarthritis Research Consortium of Canada.
† Agreement of ≥2 of 3 readers for binary variables and of 3 readers for continuous variables from wave 3.
‡ Estimated from a model in which all independent variables (time, reader, and wave) and the outcome were standardized.
§ Common reference: ASAS MRI of the SI joints.
¶ Refs. 9–11.
# Quadratic transformation led to a better model goodness of fit (quasi-likelihood under the independence model criterion).
** Refs. 3 and 12–14.

Table 3. Baseline score and standardized yearly rate of change (ROC) of structural imaging outcomes over 5 years of follow-up in patients with early axial spondyloarthritis (SpA) fulfilling the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axial SpA*

Imaging outcomes	Baseline score (range 313–344)†	Standardized ROC per year‡	Relative standardized ROC§	Relative standardized ROC per modality and anatomic site
Structural lesions (radiograph of the SI joints)¶				
mNY dichotomous, no. (%)	73 (21.2)	0.044	1	1
mNY 1-grade change#	NA	0.126	2.864	2.864
mNY 1-grade change and value ≥2**	NA	0.119	2.705	2.705
mNY continuous grade (range 0–8)	1.7 ± 1.8	0.043	0.977	0.977
Structural lesions (MRI of the SI joints)††				
≥5 fatty lesions and/or erosions, no. (%)	66 (19.5)	0.238‡‡	5.409	1
≥3 erosions, no. (%)	60 (17.7)	0.015	0.341	0.063
≥3 fatty lesions, no. (%)	56 (16.5)	0.274‡‡	6.227	1.151
No. of fatty lesions and/or erosions (range 0–80)	2.9 ± 4.9	0.111	2.523	0.466
No. of erosions (range 0–40)	1.3 ± 2.2	0.030	0.682	0.126
No. of fatty lesions (range 0–40)	1.5 ± 3.5	0.140	3.182	0.588
Total structural lesions (range 0–144)§§	3.4 ± 5.9	0.115	2.614	0.483
Total structural lesions without sclerosis (range 0–104)	3.2 ± 5.8	0.124	2.818	0.521
Structural lesions (radiograph of the spine)¶¶				
≥1 syndesmophyte, no. (%)	19 (5.5)	0.037	0.841	1
mSASSS score (range 0–72)	0.3 ± 1.3	0.043	0.977	1.162
Structural lesions (MRI of the spine)##				
≥5 fatty lesions, no. (%)	5 (1.6)	–0.013	0.295	1
Total structural lesions (range 0–322)***	0.4 ± 1.0	0.016	0.364	1.231
No. of fatty lesions (range 0–92)	0.3 ± 0.8	0.008	0.182	0.615
No. of corner erosions (range 0–92)	0.1 ± 0.2	0.012	0.273	0.923
No. of corner bone spurs (range 0–92)	0.1 ± 0.3	0.027	0.614	2.077

* Values are the mean ± SD unless indicated otherwise. mNY = modified New York criteria; MRI = magnetic resonance imaging; mSASSS = modified Stoke Ankylosing Spondylitis Spine Score; NA = not applicable; SI = sacroiliac.
† Agreement of ≥2 of 3 readers for binary variables and of 3 readers for continuous variables from wave 3.
‡ Estimated from a model in which all independent variables (time, reader, and wave) and the outcome were standardized.
§ Common reference: mNY.
¶ Refs. 15 and 16.
# Change of at least 1 grade in at least 1 SI joint.
** Change of at least 1 grade in at least 1 SI joint, but with a 5-year grade ≥2 in the worsened joint.
†† Ref. 18.
‡‡ Quadratic transformation led to a better model goodness of fit (quasi-likelihood under the independence model criterion).
§§ Fatty lesions, erosions, sclerosis, and partial ankylosis/total ankylosis.
¶¶ Ref. 19.
Refs. 20 and 21.
*** Erosions, fat infiltration, bone spurs, and ankylosis.

a standardized rate of 0.037 for ≥1 syndesmophyte and of 0.043 for the continuous mSASSS.

DISCUSSION

In this prospective observational study, we have shown that in patients with early axial SpA, MRI outcomes of inflammation are more sensitive to change in the SI joints than in the spine. In addition, pelvic radiographs yield low sensitivity to change in detecting structural damage, while fatty lesions detected on MRI-SI joints emerges as a promising alternative. In contrast, MRI-spine is not better than radiography of the spine in detecting structural changes in patients with early axial SpA.

In the current study, we directly compared, for the first time, inflammation outcomes on MRI-SI joints and MRI-spine and have shown that the former are more sensitive to change. Inflammation on MRI-spine remained low and essentially unchanged over a period of 5 years. Different from previous studies evaluating the sensitivity to change of imaging outcomes over shorter periods, we have applied an analytical technique (integrated analysis) that we have previously shown to be robust for the evaluation of change over long periods of follow-up, especially with outcomes that are expected to occur infrequently over time (8). Of note, combination algorithms (e.g., agreement between 2 of 3 readers) are not needed when using this method. Instead, each individual reader score is analyzed as it is in an assumption-free manner that, to some extent, handles across-reader variability.

The ASAS/OMERACT MRI working group has previously compared different (continuous) scores to quantify inflammation on MRI-SI joints (22). In a multireader exercise, the SPARCC method has been shown to be the most reliable and sensitive to change among patients with radiographic axial SpA. The current

study adds to these data by showing that both the continuous SPARCC score and the binary ASAS definition of a positive MRI-SI joint yield good sensitivity to change in the entire spectrum of axial SpA (including nonradiographic axial SpA) during the early phases of the disease.

The same group performed a similar exercise for MRI-spine (also in radiographic axial SpA) (3). This experiment has shown discrepant reliability results for the comparison between the 6-discovertebral unit (DVU) SPARCC score, the Ankylosing Spondylitis Spine MRI Activity score, and the Berlin method (SPARCC performed better when using the intraclass correlation coefficient but worse when using the smallest detectable change). All methods yielded excellent sensitivity to change according to Guyatt's effect size. Here, we compared the 23-DVU SPARCC score to the Berlin method and 2 binary outcomes and found that all yield very poor sensitivity to change. Of note, these studies differ in several aspects, including the reading methods and population. In fact, our early axial SpA population had lower baseline levels of inflammation compared to that in patients from the ASAS/OMERACT exercise (mean $\pm$ SD Berlin score 0.9 ± 2.7 versus 6 ± 9.0, respectively), which may hinder the detection of change, which we have shown before to be small in early axial SpA (17). Of note, in patients with nonradiographic axial SpA and high disease activity selected for randomized controlled trials, inflammation on MRI-spine performed well both in terms of sensitivity to change and in discriminating response between treatment arms (23,24). This confirms that the ability of the scoring methods to detect change is not only dependent on their intrinsic characteristics, but also on the population in which they are applied.

A recent study, also from the DESIR cohort, has shown that net progression from mNY negative to mNY positive (i.e., considering measurement error) is very limited (16). In the current study, we have additionally shown that the change in the mNY (continuous) grading is as poorly sensitive to change as the mNY binary score (relative rate of ~1). However, the change of at least 1 grade in at least 1 SI joint, with or without considering the change between grade 0 and grade 1, performs better in detecting change (16,25).

Information on the sensitivity to change of MRI-SI joints structural outcomes is very scarce (26). To the best of our knowledge, no previous formal comparison with radiography of the SI joints scores has been performed thus far. We have found that $\geq$3 fatty lesions on MRI-SI joints largely outperform all radiography of the SI joints outcomes. Erosions, however, performed poorly in this early population. Thus, our study yields encouraging data supporting MRI (in particular fatty lesions) as an alternative to radiographs in detecting change of structural damage at the SI joints. In contrast, in the spine, we found no evidence that MRI is better than radiographs in detecting change of structural damage. Despite the disappointing results with MRI, our results are in line with previous studies, showing that spinal radiographic progression can be detected even in early phases of the disease (4,27). A recent study has shown that low-dose computerized tomography of the spine is more sensitive at detecting new syndesmophytes than conventional radiographs, which promises to further expand our ability to detect change in axial damage (28).

Our study has some limitations. First, not all available scoring systems were assessed. However, to the best of our knowledge, this is, so far, the largest direct comparison across scores, which includes those currently more often used in research and clinical practice. Second, we did not assess all domains of the OMERACT filter, namely validity, reliability, and feasibility (2). Thus, we cannot, and do not claim to, evoke superiority of one score over others based on our data alone. Instead, our results should be interpreted in light of the literature already informing on these aspects but falling short on direct comparisons of sensitivity to change. Third, the observed levels of inflammation, structural damage, and changes over time are limited in this cohort, especially in the spine, which reduces the possibility of detecting differences across methods. Finally, our data are limited to patients with early axial SpA; thus, our findings cannot be generalized to all patients with axial SpA from clinical practice, especially those with more advanced disease (i.e., with radiographic axial SpA).

In conclusion, we have shown that MRI inflammation scores are more sensitive to change in the SI joints than in the spine. Also, radiography of the SI joints structural outcomes are less sensitive to change compared to fatty lesions on MRI-SI joints. In contrast, MRI-spine is no better than radiography of the spine in detecting structural changes in this early axial SpA cohort. These data may help in prioritizing imaging scoring methods in subsequent observational or interventional studies in early axial SpA.

ACKNOWLEDGMENTS

The authors thank the different regional participating centers: Maxime Dougados (Paris-Cochin B), Andre Kahan (Paris-Cochin A), Philippe Dieudé (Paris-Bichat), Bruno Fautrel (Paris-La Pitie-Salpetriere), Francis Berenbaum (Paris-Saint-Antoine), Pascal Claudepierre (Creteil), Maxime Breban (Boulogne-Billancourt), Bernadette Saint-Marcoux (Aulnay-sous-Bois), Philippe Goupille (Tours), Jean Francis Maillefert (Dijon), Emmanuelle Dernis (Le Mans), Daniel Wendling (Besancon), Bernard Combe (Montpellier), Liana Euller-Ziegler (Nice), Pascal Richette (Paris Lariboisiere), Pierre Lafforgue (Marseille), Patrick Boumier (Amiens), Martin Soubrier (Clermont Ferrand), Nadia Mehsen (Bordeaux), Damien Loeuille (Nancy), Rene-Marc Flipo (Lille), Alain Saraux (Brest), Xavier Mariette (LeKremlin-Bicetre), Alain Cantagrel (Toulouse), and Olivier Vittecoq (Rouen). They also thank the research nurses, the staff members of the Clinical Research Unit of Paris Centre, the staff members of the Biological Resource Center of Bichat Hospital, the staff members of the Department of Statistics of Nîmes, and all the investigators, and in particular Jerome Allain, Emmanuelle Dernis, Salah Ferkal, Clement Prati, Marie-Agnes Timsit, and Eric Toussirot, for active patient recruitment and monitoring. The authors are grateful to Anna Moltó and Miranda van Lunteren for their work in preparing the database used in this study; and to all DESIR readers, without whom this study would not be possible.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final

version to be submitted for publication. Dr. Sepriano had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Sepriano, Ramiro, van der Heijde, Dougados, Landewé.

Acquisition of data. Claudepierre, Feydy, Reijnierse, Loeuille.

Analysis and interpretation of data. Sepriano, Ramiro, van der Heijde, Dougados, Landewé.

REFERENCES

1. Mandl P, Navarro-Compan V, Terslev L, Aegerter P, van der Heijde D, D'Agostino MA, et al. EULAR recommendations for the use of imaging in the diagnosis and management of spondyloarthritis in clinical practice. Ann Rheum Dis 2015;74:1327–39.
2. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745–53.
3. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Ostergaard M, et al. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. J Rheumatol 2007;34:862–70.
4. Ramiro S, van Tubergen A, Stolwijk C, Landewé R, van de Bosch F, Dougados M, et al. Scoring radiographic progression in ankylosing spondylitis: should we use the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) or the Radiographic Ankylosing Spondylitis Spinal Score (RASSS)? Arthritis Res Ther 2013;15:R14.
5. Wanders AJ, Landewé RB, Spoorenberg A, Dougados M, van der Linden S, Mielants H, et al. What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 2004;50:2622–32.
6. Ramiro S, Claudepierre P, Sepriano A, van Lunteren M, Molto A, Feydy A, et al. Which scoring method depicts spinal radiographic damage in early axial spondyloarthritis best? Five-year results from the DESIR cohort. Rheumatology (Oxford) 2018;57:1991–2000.
7. Dougados M, d'Agostino MA, Benessiano J, Berenbaum F, Breban M, Claudepierre P, et al. The DESIR cohort: a 10-year follow-up of early inflammatory back pain in France: study design and baseline characteristics of the 708 recruited patients. Joint Bone Spine 2011;78:598–603.
8. Sepriano A, Ramiro S, van der Heijde D, Dougados M, Claudepierre P, Feydy A, et al. Integrated longitudinal analysis does not compromise precision and reduces bias in the study of imaging outcomes: a comparative 5-year analysis in the DESIR cohort. Semin Arthritis Rheum 2020;50:1394–9.
9. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703–9.
10. Rudwaleit M, Jurik AG, Hermann KG, Landewé R, van der Heijde D, Baraliakos X, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009;68:1520–7.
11. Lambert RG, Bakker PA, van der Heijde D, Weber U, Rudwaleit M, Hermann KG, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 2016;75:1958–63.
12. Hermann KG, Baraliakos X, van der Heijde DM, Jurik AG, Landewé R, Marzo-Ortega H, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. Ann Rheum Dis 2012;71:1278–88.
13. De Hooge M, van den Berg R, Navarro-Compan V, Reijnierse M, van Gaalen F, Fagerli K, et al. Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? Ann Rheum Dis 2016;75:1308–14.
14. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Krishnananthan R, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:502–9.
15. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
16. Dougados M, Sepriano A, Molto A, van Lunteren M, Ramiro S, de Hooge M, et al. Sacroiliac radiographic progression in recent onset axial spondyloarthritis: the 5-year data of the DESIR cohort. Ann Rheum Dis 2017;76:1823–8.
17. Madari Q, Sepriano A, Ramiro S, Molto A, Claudepierre P, Wendling D, et al. 5-year follow-up of spinal and sacroiliac MRI abnormalities in early axial spondyloarthritis: data from the DESIR cohort. RMD Open 2020;6:e001093.
18. Weber U, Lambert RG, Østergaard M, Hodler J, Pedersen SJ, Maksymowych WP. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. Arthritis Rheum 2010;62:3048–58.
19. 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;64:127–9.
20. Ostergaard M, Maksymowych WP, Pedersen SJ, Chiowchanwisawakit P, Lambert RG. Structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis: definitions, assessment system, and reference image set. J Rheumatol 2009;36 Suppl 84:18–34.
21. Krabbe S, Sorensen IJ, Jensen B, Moller JM, Balding L, Madsen OR, et al. Inflammatory and structural changes in vertebral bodies and posterior elements of the spine in axial spondyloarthritis: construct validity, responsiveness and discriminatory ability of the anatomy-based CANDEN scoring system in a randomised placebo-controlled trial. RMD Open 2018;4:e000624.
22. Van der Heijde DM, Landewé RB, Hermann KG, Jurik AG, Maksymowych WP, Rudwaleit M, et al. Application of the OMERACT filter to scoring methods for magnetic resonance imaging of the sacroiliac joints and the spine: recommendations for a research agenda at OMERACT 7. J Rheumatol 2005;32:2042–7.
23. Maksymowych WP, Dougados M, van der Heijde D, Sieper J, Braun J, Citera G, et al. Clinical and MRI responses to etanercept in early non-radiographic axial spondyloarthritis: 48-week results from the EMBARK study. Ann Rheum Dis 2016;75:1328–35.
24. Sieper J, van der Heijde D, Dougados M, Mease PJ, Maksymowych WP, Brown MA, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72:815–22.
25. Poddubnyy D, Rudwaleit M, Haibel H, Listing J, Marker-Hermann E, Zeidler H, et al. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. Ann Rheum Dis 2011;70:1369–74.
26. Baraliakos X, van der Heijde D, Braun J, Landewé RB. OMERACT magnetic resonance imaging initiative on structural and inflammatory lesions in ankylosing spondylitis: report of a special interest group at OMERACT 10 on sacroiliac joint and spine lesions. J Rheumatol 2011;38:2051–4.

27. Cho SK, Sakai R, Nanki T, Koike R, Watanabe K, Yamazaki H, et al. A comparison of incidence and risk factors for serious adverse events in rheumatoid arthritis patients with etanercept or adalimumab in Korea and Japan. Mod Rheumatol 2014;24: 572–9.

28. De Koning A, de Bruin F, van den Berg R, Ramiro S, Baraliakos X, Braun J, et al. Low-dose CT detects more progression of bone formation in comparison to conventional radiography in patients with ankylosing spondylitis: results from the SIAS cohort. Ann Rheum Dis 2018;77:293–9.