Treating spondyloarthritis early: does it matter? Results from a systematic literature review

Capelusnik, D.; Benavent, D.; Heijde, D. van der; Landewe, R.; Poddubnyy, D.; Tubergen, A. van; ... ; Ramiro, S.

Rheumatology, 2023, **62**, 1398–1409
https://doi.org/10.1093/rheumatology/keac532
Advance access publication 13 September 2022
Systematic review and meta analysis

Clinical science

Treating spondyloarthritis early: does it matter? Results from a systematic literature review

Dafne Capelusnik [1,2,*], Diego Benavent [3], Désirée van der Heijde [4], Robert Landewé[5,6], Denis Poddubnyy [7], Astrid van Tubergen [8], Louise Falzon [9], Victoria Navarro-Compán [3,†], Sofia Ramiro [4,6,†]

[1]Department of Medicine, Care and Public Health Research Institute CAPHRI, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands
[2]Department of Rheumatology, Instituto de Rehabilitación Psicofísica (IREP), Ciudad de Buenos Aires, Argentina
[3]Department of Rheumatology, Hospital Universitario La Paz, IdiPaz, Madrid, Spain
[4]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[5]Department of Clinical Immunology and Rheumatology, Amsterdam University Medical Center, Amsterdam, The Netherlands
[6]Department of Rheumatology, Zuyderland Medical Center, Heerlen, The Netherlands
[7]Department of Gastroenterology, Infectiology and Rheumatology, Charité Universitätsmedizin Berlin, Berlin, Germany
[8]Department of Medicine, Division of Rheumatology, Maastricht University Medical Center, and School for Public Health and Primary Care (CAPHRI), Maastricht University, Maastricht, The Netherlands
[9]University of Sheffield, Sheffield, UK

*Correspondence to: Dafne Capelusnik, Department of Medicine, Care and Public Health Research Institute CAPHRI, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands. E-mail: capelusnikdafne@gmail.com

[†]Victoria Navarro-Compán and Sofia Ramiro contributed equally to this study.

Abstract

Objective: To summarize evidence on the relationship between early treatment (definition based on symptom/disease duration or radiographic damage) and treatment clinical response in patients with SpA.

Methods: A systematic literature review was conducted in studies on SpA patients treated with NSAIDs or biological/targeted synthetic DMARDs addressing the impact of symptom/disease duration or presence of radiographic damage on treatment response assessed by any disease activity outcome. For categorical outcomes, relative risk, relative risk ratio and number needed to treat were calculated, and for continuous outcomes, differences in differences, to compare groups stratified based on symptom/disease duration or the presence of radiographic damage.

Results: From the 8769 articles retrieved, 25 were included and 2 added by hand-search, all in axial SpA (axSpA), most of them with low risk of bias. Twenty-one studies compared groups based on symptom duration ($n=6$) or disease duration ($n=15$) and seven studies based on absence/presence of radiographic damage (two studies used two comparisons). When early axSpA was defined by symptom duration (<5 years) in randomized controlled trials, early treatment was associated with better outcomes in patients with non-radiographic axSpA [$n=2$, ASAS40 relative risk ratio 5.24 (95% CI 1.12, 24.41) and 1.52 (0.60, 3.87)] but not in radiographic axSpA ($n=1$) [ASAS20 0.96 (0.53–1.73)]. When early axSpA was defined based on disease duration or radiographic damage, no differences were found between groups.

Conclusion: Evidence towards better outcomes in early axSpA is very limited and restricted to non-radiographic axSpA and <5 years symptom duration. When early axSpA is defined based on disease duration or radiographic damage, no differences in response to treatment are found.

Keywords: axial spondyloarthritis, early disease, treatment, response, outcomes, systematic review

Rheumatology key messages

- Evidence on better response to treatment in early axial SpA (axSpA) is very limited.
- Better treatment outcomes are found in early non-radiographic axSpA and <5 years of symptom duration.
- Well-designed analyses are needed to claim an effect of symptom duration on treatment response in axSpA.

Introduction

The concept of a 'therapeutic window of opportunity' refers to the period of time where the initiation of treatment increases the potential of improving long-term outcomes. This, together with the treat to target strategy, has been extensively explored in RA [1, 2]. Although such a concept has not been formally investigated in SpA, a similar belief can be found in the literature, where early treatment has been interpreted as being more

Received: 16 June 2022. **Accepted:** 30 August 2022

beneficial [3, 4]. Most of the evidence supporting this idea comes from randomized controlled trial (RCT) *post hoc* analyses comparing the effect of biological DMARDs (bDMARDs) *vs* placebo in patients with axial SpA (axSpA) [5]. However, some misconceptions may arise based on this assumption: those studies were designed to assess differences in response between active treatment—bDMARDs—and placebo, and there is no formal comparison of the treatment effect between the two stages of the disease (early and established disease), taking both interventions into account. What we are really interested in is whether the effect of treatment is different in patients with early *vs* established disease, which means that both interventions (i.e. bDMARD and placebo) need to be considered in the analysis to allow for proper conclusions.

To understand what 'early treatment' implies, it is important to clarify what 'early disease' means. This is what sparked, under the auspices of Assessment of SpondyloArthtitis international Society (ASAS), the rationale behind the ASAS-SPEAR (Spondylarthritis EARly definition) project, aiming to develop a consensual definition of 'early spondylarthritis'. As an initial step, a systematic literature review (SLR) was performed to review and summarize the existing terms used to define early SpA. This revealed that a definition of early disease was mainly used to refer to the axSpA subtype and recognized substantial heterogeneity among the identified definitions, mainly comprising time-based definitions with different cut-off points of symptoms or disease duration, as well as definitions based on the absence of radiographic damage [non-radiographic axSpA (nr-axSpA)] [6].

In order to further inform ASAS members, the current SLR was performed to summarize the existing evidence of the relationship between early treatment (based on definitions of symptom duration, disease duration or the absence of radiographic damage) and clinical response in patients with SpA treated with NSAIDs, bDMARDs or targeted synthetic DMARDs (tsDMARDs).

Methods

Search strategy and studies selection

The protocol of the SLR was developed by the ASAS-SPEAR Steering Committee and subsequently registered in PROSPERO (CRD42020173571). This review is reported with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocol (PRISMA-P) guidelines.

The search was led by an expert librarian (L.F.) using Medline, EMBASE and the Cochrane Library, until 28 April 2021 (Supplementary Data S1, available at *Rheumatology* online), with a search strategy appropriately adapted for each database, without language restriction. Eligible ACR and EULAR abstracts from 2019–20 congresses were obtained from the corresponding websites. In addition, since data were sometimes not published in scientific journals but included in the information submitted by pharmaceutical companies to regulatory agencies, hand searches were conducted for each drug, in the Food and Drug Administration (FDA) and European Medicines Agency (EMA) websites.

The research question was structured according to the PICO framework (Population, Intervention, Comparator, Outcome) [7]. The eligible study types were RCTs and cohort studies in patients with SpA addressing the impact of symptom duration or disease duration, or the presence of radiographic damage on treatment (NSAIDs or b/tsDMARDs) response. Based on a cut-off of symptom/disease duration or the absence/presence of radiographic damage, treatment response in the groups of 'early' and 'established' disease was compared. For the RCTs the Population was defined as adult patients (>18 years), with SpA, either axSpA or peripheral SpA (pSpA), starting treatment with NSAIDs, bDMARDs or tsDMARDs, or placebo; the Intervention was treatment in early disease based on any definition (symptom duration, disease duration or the absence of radiographic structural damage); the Comparator was treatment in established disease (meaning 'not early disease'); and the Outcome was treatment response according to (i) disease activity [ASAS response criteria (ASAS20, ASAS 40, ASAS5/6), ASAS partial remission (ASAS-PR) [8, 9], Ankylosing Spondylitis Disease Activity Score Major Improvement (ASDAS-MI), ASDAS Clinically Important Improvement (ASDAS-CII), ASDAS Low Disease Activity (ASDAS-LDA), ASDAS Inactive Disease (ASDAS-ID) [10], and BASDAI 50 response (BASDAI50) [11], change since baseline of continuous outcomes, such as ASDAS [12], BASDAI [13], CRP, Patient Global Assessment]; (ii) physical function in Bath AS Functional Index (BASFI) [14]; (iii) spinal mobility—Bath AS Metrology Index (BASMI) [15]; and (iv) quality of life or overall functioning and health [improvement in Short Form-36 (SF-36) [16] and ASAS Health Index (ASAS HI)] [17]. Only RCTs assessing the outcome at the timing of the primary endpoint were included in order to allow for the comparison between the intervention and the comparator arm, i.e. the placebo arm. For observational studies, Population was the same; Intervention was the predictive role of symptom duration, disease duration or the absence of radiographic structural damage on treatment response; in these studies, there was no Comparator; and the Outcome was the effect size of the predictive treatment response measured by disease activity, function and quality of life or overall functioning and health as previously mentioned. Cohort studies were only included if the predictive role of symptom/disease duration or the absence of radiographic damage was analysed in multivariable analyses adjusted for potential confounders, or if an interaction between treatment and early/established disease on treatment outcomes was reported with results stratified in the two groups.

Of note, we were mainly interested in the effect of symptom duration and not on that of disease duration, given that disease duration is substantially influenced by the known diagnostic delay in SpA, especially in axSpA [18]. Nevertheless, for a comprehensive understanding of the effect of time, we also added the term 'disease duration' to the SLR.

Data extraction and risk of bias assessment

The screening of the studies was performed through title and abstract by two reviewers independently (D.C. and D.B.) using Rayyan [19]. The abstracts considered relevant were selected for full text assessment.

The data extracted included type of study (RCTs or cohort study), population [axSpA, radiographic axSpA (r-axSpA and nr-axSpA) or pSpA], definition of early SpA used (symptom duration or disease duration and corresponding cut-off used or the absence of radiographic structural damage in the sacroiliac joints according to the modified New York criteria [20]), number of patients included, age, gender, symptom duration, disease duration, therapeutic interventions (including comparator) and treatment clinical outcomes. Both reviewers independently identified eligible studies and extracted the

data, including the risk of bias (RoB) assessment according to the Cochrane Collaboration's tool for RCTs, Newcastle-Ottawa Scale for case–control studies and the 'Hayden tool' for observational studies [21–23]. Disagreements were discussed until consensus was achieved, and whenever necessary, two methodologists (S.R. and V.N.-C.) were involved.

Data analysis

A descriptive analysis was performed stratified according to the definition of early SpA used and according to the study type. For categorical outcomes in RCTs, the following three measures were calculated: (i) number needed to treat (NNT) [24], which is the number of patients we need to treat in order to achieve one additional patient reaching the outcome, (ii) relative risk (RR) [25], which represents the likelihood of achieving the outcome in the active treatment group *vs* the placebo group, and (iii) relative risk ratio (RRR) [26], which is the ratio between the RR of active treatment (e.g. bDMARD) in early disease *vs* the RR in established disease. The RRR allows assessment of the effect of active treatment—e.g. bDMARD—*vs* placebo in early *vs* established disease. An RRR >1 reflects a higher likelihood of achieving the outcome in patients with early disease treated with bDMARD compared with established disease; in other words, the treatment effect is higher in early disease compared with established disease. A 95% CI around the RRR was also calculated to provide the uncertainty around the estimate. As with the RRR a formal comparison is made between both groups (early and established disease); this method, together with the 95% CI, prevailed for the interpretation of the results. For the continuous variables, the difference in differences [27] was calculated, which is the difference between improvement since baseline (treatment effect) in early and established disease. Lastly, from the cohort studies, the predictive effect of the outcomes of interest, expressed as odds ratio or coefficient B—with corresponding 95% CI—was extracted from the corresponding multivariable model.

If data from at least two RCTs with the same definition of early disease and assessing the same outcome at the same timepoint were available, a pooled analysis was performed. Pooled RRR were estimated using a random-effects model, chosen to be conservative, independently of the statistical heterogeneity. Microsoft Excel and R-Cran V.3.5.1 software with the package 'data.table' were used for the statistical analysis.

Results

A total of 8769 articles were identified with the search. After removal of duplicates, 6553 articles were screened by title and abstract and 492 were selected for full text assessment. A total of 26 studies were included for data extraction: 25 from the bibliographic search (20 full-text articles [5, 28–46] and 5 congress abstracts [47–51]) and 2 sub-analyses [52, 53] were added by 'hand search' extracted from the FDA website (supplementary Fig. S1, available at *Rheumatology* online), one of the latter representing additional information from an already included study [52]. Notably, all included studies were on axSpA and active treatment was bDMARDs compared with placebo.

Fig. 1 illustrates the included studies stratified according to the criteria used to define early axSpA, namely time criteria (symptom or disease duration, $n = 6$ [28, 29, 32, 33, 47, 52] and $n = 15$ studies [5, 30, 31, 34–40, 48–51, 53], respectively) or radiographic criterion (absence *vs* presence of radiographic damage, i.e. r-axSpA *vs* nr-axSpA, $n = 7$ studies [33, 41–46]), and according to the type of study (RCTs or cohort study). Two of the included studies [33, 44, 52] addressed treatment outcomes based on time and radiographic criteria, and could therefore be included for both aspects. Most studies were at low RoB. A more detailed description of the included studies, including name of the trial or cohort, demographic characteristics intervention and comparator, and RoB can be found in supplementary Tables S1–S4, available at *Rheumatology* online.

Studies comparing early vs established axSpA based on symptom duration definition

From the six studies with early disease based on symptom duration, four were RCTs, all of them using the cut-off of 5 years and thus comparing <5 years *vs* ≥5 years, and all of them using TNF inhibitors (TNFi) (adalimumab and certolizumab pegol); three of the studies were on nr-axSpA and one on axSpA, including patients with r-axSpA and nr-axSpA. In patients with nr-axSpA, like in Sieper *et al.* (2012) [28], a higher percentage of patients with early disease achieving ASAS40 was found in the bDMARDs arm (48% *vs* 31% established disease) but not in the placebo arm (6% *vs* 20%); hence, in our analysis, bDMARD treatment in early disease was associated with a higher RR for ASAS40 (8.2 *vs* 1.6). This was also reflected in a higher RRR [5.24 (95% CI 1.12, 24.41)], meaning that patients with early disease had a 5

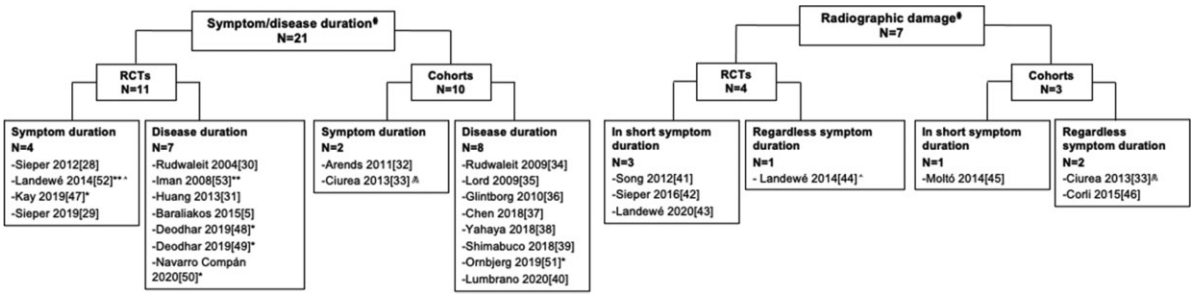

Figure 1. Summary of included studies according to definition of early or established disease based on symptom/disease duration or radiographic damage. #One of the 27 included studies reported treatment outcomes according to both definitions of early *vs* established disease, i.e. disease duration and radiographic damage. *Congress abstract. **Randomized controlled trials (RCTs) extracted in a second stage from the Food and Drug Administration website. ^Study addressing treatment outcome based on symptom duration and radiographic damage definition, and therefore included separately for both aspects. ^^Study addressing treatment outcome based on disease duration and radiographic damage definition, and therefore included separately for both aspects. References and details of the included studies in supplementary Table S1, available at *Rheumatology* online

times higher likelihood of achieving ASAS40 compared with established disease when comparing bDMARDs with placebo. The NNT was lower for early compared with established disease (2.4 *vs* 9.1). In Kay *et al.* (2019) [47], also in nr-axSpA, higher percentages of patients with early disease achieved ASAS40 in both arms (59% *vs* 37% early *vs* established disease for bDMARD and 12% *vs* 11% for placebo, respectively) [72]. However, although the RR and the NNT were in favour of early treatment compared with the group with established disease (5.0 *vs* 3.3 and 2.1 *vs* 3.9, respectively), when assessing the RRR [1.52 (95% CI 0.60, 3.87)], these differences did not reach statistical significance. When pooling the data from the only two studies assessing ASAS40 in nr-axSpA [28, 47], an RRR of 3.74 (95% CI 1.47, 9.56) was obtained, still favouring the group with early disease. The third RCT in nr-axSpA, Sieper *et al.* (2019) [29], showed that patients achieving ASDAS-ID and ASAS-PR had shorter symptom duration than those not achieving these outcomes. However, no response outcomes were given per group stratified according to symptom duration, hampering comparisons. Finally, in Landewé *et al.* (2014) [52], no difference in treatment response was observed between early and established disease according to symptom duration [ASAS20 RRR 0.96 (95% CI 0.53, 1.73)] (Table 1). In the cohort studies ($n = 2$), when exploring symptom duration as a predictor of response, no significant effect was found [32, 33].

Studies comparing early vs established axSpA based on disease duration definition

When early axSpA was defined based on disease duration in RCTs ($n = 7$), only one study included patients with nr-axSpA: Navarro-Compán *et al.* (2020) showed higher RR and a lower NNT for ASAS40 in patients with early disease, but the RRR comparing the treatment effect (*vs* placebo) between disease subgroups showed no significant differences (Table 2) [50].

In almost all the remaining studies, which included patients with r-axSpA, there was a higher proportion of patients with early *vs* established disease achieving the outcomes in each of the bDMARDs and the placebo arms. However, when formally comparing them, the RRR showed no differences between the groups [5, 30, 31, 48, 49, 53]. As for continuous outcomes, all comparisons come from Baraliakos *et al.* (2015) [5], in which a numerically higher change from baseline was achieved in established disease, except when comparing change of Patient Global Assessment in patients with <2 years *vs* ≥5–10 and >10 years, and CRP in patients with <2 years *vs* ≥5–10 years, where a higher change was achieved in early disease (Table 2) [5].

In cohort studies, when using disease duration as a predictor of treatment response ($n = 8$), Chen *et al.* (2018) [37] showed that shorter disease duration (≤6 years) was an independent predictor of achieving minimal clinically important difference in the Physical Component Summary of SF-36. In Lubrano *et al.* (2020) [40], shorter disease duration was associated with a higher likelihood of BASMI and BASFI improvement (Δ 1 and 2 points, respectively). In the remaining cohort studies, early disease was not associated with better response to treatment (supplementary Table S5, available at *Rheumatology* online) [30, 35, 36, 38, 39, 51].

Studies comparing early vs established axSpA based on radiographic damage

Next, when early disease was defined based on radiographic damage (nr-axSpA *vs* r-axSpA) in RCTs with patients with <5 years of disease duration ($n = 3$), the proportion of patients achieving the outcome in the active arm varied across the different measures, some favouring r-axSpA and others nr-axSpA [41–43]. In the only study with available data also in the placebo arm, Sieper *et al.* (2016) [42], the calculated RR and NNTs favoured r-axSpA, but the RRR showed no difference between the two groups. Lastly, in Landewé *et al.* (2014) [44] comparing nr-axSpA *vs* r-axSpA in patients without any restriction of symptom or disease duration, results were again inconsistent: in the raw data extracted from the study, the majority of the outcomes were more frequently achieved by patients with early disease (except for ASAS-5/6), but RR, NNTs and RRR were inconsistent with each other; notwithstanding RRR reflected no significant difference between the groups (Tables 3 and 4).

In cohort studies ($n = 3$), the absence of radiographic damage (nr-axSpA) was not associated with better response to treatment compared with its presence (i.e. r-axSpA) (supplementary Table S6, available at *Rheumatology* online) [33, 45, 46].

Discussion

In this SLR, aimed at summarizing the evidence on the relationship between early treatment and treatment clinical response in patients with SpA, we found that when defining early disease based on symptom duration, combined results from two studies in nr-axSpA showed that treatment with bDMARDs may lead to better outcomes compared with established nr-axSpA. However, these differences were not found in axSpA (studies including nr- and r-axSpA together), or when defining early disease based on disease duration or radiographic damage. Additionally, no data on pSpA or SpA (as a whole, including axSpA and pSpA) were found.

To the best of our knowledge, this is the first SLR addressing the comparison of treatment effect in early *vs* established disease based on a time-based definition; however, we did not find robust evidence to conclude that there is a benefit of early treatment. In fact, there was only scarce evidence pointing towards such a beneficial effect when early disease is defined based on symptom duration in nr-axSpA, and still not free of some inconsistencies. Pooled data from two studies suggest that early nr-axSpA treated with bDMARDs has almost 4 times higher likelihood of achieving ASAS40 compared with established disease, meaning that the effect of bDMARDs *vs* placebo is 4 times higher in early *vs* established disease. Still, it is important to note that the result of this pooled data (meta-analysis) carries more weight from one RCT, namely Sieper *et al.* (2019) [28]. In the *post hoc* analysis from Kay *et al.* (2019) [47], the other study included in the meta-analysis, despite the numerically higher responses in early disease supporting the results, there was no statistically significant difference between the groups. Results of this SLR need to be interpreted cautiously. Studies reporting the *post hoc* analyses we were interested in are scarce, and the definition used for early treatment among these was heterogeneous. Thus, current results must be interpreted in the context of a possible publication bias, as negative studies are less often published.

Table 1. Assessment of treatment response in randomized controlled trials comparing early *vs* established disease defined based on symptom duration

Study	Time point	Drug	Population	Symptom duration (years)	Outcome drug group		outcome placebo group		*P*-value	RR (early *vs* established)	RRR (95% CI)	NNTs (early *vs* established)
					Early disease	Established disease	Early disease	Established disease				
ASAS20												
Landewé 2014[a] [52]	w12	CZP	axSpA	<5 *vs* ≥5	28/50 (56%)	36/61 (59%)	28/74 (38%)	25/65 (38%)	0.98[b]	1.5 *vs* 1.5	0.96 (0.53, 1.73)	5.5 *vs* 4.8
ASAS40												
Sieper 2013 [28]	w12	ADA	nr-axSpA	<5 *vs* ≥5	16/33 (48%)	17/55 (31%)	2/34 (6%)	11/56 (20%)	0.02[b]	8.2 *vs* 1.6	5.24 (1.12, 24.41)	2.4 *vs* 9.1
Kay 2019[c] [47]	w12	CZP	nr-axSpA	<5 *vs* ≥5	47/80 (59%)	29/79 (37%)	9/77 (12%)	9/81 (11%)	NR	5.0 *vs* 3.3	1.52 (0.60, 3.87)	2.1 *vs* 3.9
	w52				52/80 (65%)	38/79 (48%)	14/77 (18%)	11/81 (14%)	NR	3.6 *vs* 3.5	1.01 (0.46, 2.20)	2.1 *vs* 2.9
ASDAS-MI												
Kay 2019[c] [47]	w12	CZP	nr-axSpA	<5 *vs* ≥5	37/80 (46%)	19/79 (24%)	7/77 (9%)	3/81 (4%)	NR	5.1 *vs* 6.5	0.78 (0.19, 3.16)	2.7 *vs* 4.9
	w52				44/80 (55%)	31/79 (39%)	6/77 (8%)	5/81 (6%)	NR	7.1 *vs* 6.4	1.11 (0.34, 3.66)	2.1 *vs* 3.0

Study	Time point	Drug	Population	Symptom duration (years)	Outcome drug group		Outcome placebo group		*P*-value	Change since baseline		
					Early disease	Established disease	Early disease	Established disease		Early	Established	DiD
Change in BASDAI												
Kay 2019[c] [47]	w12	CZP	nr-axSpA	<5 *vs* ≥5	−3.3	−2.5	−1.1	−1.1	NR	-2.2	-1.4	-1.2
	w52				−3.9	−3.3	−1.4	−1.3	NR	-2.5	-2.0	-0.5

Study	Time point	Drug	Population	Symptoms duration (years)		Disease duration (years)		*P*-value
				ASDAS-ID responders	ASDAS-ID no responders	ASDAS-ID responders	ASDAS-ID no responders	
Sieper 2019 [29]	w12	ADA	nr-axSpA	6.1 (6.2)	8.3 (8.1)			<0.001
						1.7 (2.9)	1.8 (3.6)	NS
				ASAS-PR responders	ASAS-PR no responders	ASAS-PR responders	ASAS-PR no responders	*P*-value
				5.3 (5.7)	8.0 (7.8)			<0.001
						1.7 (3)	1.8 (3.2)	NS

Cell colours: green: in favour of early disease; red: in favour of established disease; yellow: not statistically significant. Colour version available online.

[a] Data extracted in a second stage from the FDA website.

[b] Interaction *P*-value.

[c] Congress abstract.

RR: relative risk; RRR: relative risk ratio; NNTs: number of patients needed to treat; change since baseline, negative values correspond to outcome improvement; DiD: differences in differences, positive values in favour of late disease; ASAS20/40: ASAS response criteria; ASDAS-MI: Ankylosing Spondylitis Disease Activity Score Major Improvement criteria; ASDAS-ID: ASDAS Inactive Disease; ASAS-PR: ASAS partial remission; ADA: adalimumab 40 mg Q2W; CZP: certolizumab pegol 200 mg Q2W; nr-axSpA: non-radiographic SpA; NS: not significant; FDA: Food and Drug Administration.

Table 2. Assessment of treatment response in randomized controlled trials comparing early *vs* established disease defined based on disease duration

Categorical outcomes												
Study	Time point	Drug	Population	Disease duration (years)	Outcome drug group		Outcome placebo group		*P*-value	RR (early vs established)	RRR (95% CI)	NNTs (early vs established)
					Early disease	Established disease	Early disease	Established disease				
ASAS20												
Inman 2008[a] [53]	w14	GLM	r-axSpA	≤5.6 *vs* >5.6	90/148 (61%)	76/130 (58%)	9/33 (27%)	8/45 (18%)	NS	2.2 *vs* 3.3	0.68 (0.29, 1.60)	2.9 *vs* 2.5
Huang 2014 [31]	w12	ADA	r-axSpA	<2 *vs* ≥2	96/134 (72%)	58/95 (61%)	22/59 (37%)	13/56 (23%)	0.69	1.9 *vs* 2.6	0.73 (0.40, 1.35)	2.9 *vs* 2.6
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	171/225 (76%)	99/140 (71%)	22/50 (44%)	4/28 (14%)	≤0.05	1.7 *vs* 5.0	0.35 (0.13, 0.92)	3.1 *vs* 1.8
				<2 *vs* ≥5–10		108/151 (72%)		14/46 (30%)		1.7 *vs* 2.4	0.73 (0.42, 1.28)	3.1 *vs* 2.4
				<2 *vs* >10		191/289 (66%)		26/96 (27%)		1.7 *vs* 2.4	0.71 (0.44, 1.13)	3.1 *vs* 2.6
ASAS40												
Deodhar 2019[c] [49]	w16	SEC	r-axSpA	Q1 *vs* Q2	79/188 (42%)	71/188 (38%)	102/389 (26%)	78/389 (20%)	NR	1.6 *vs* 1.9	0.85 (0.59, 1.22)	6.2 *vs* 5.6
				Q1 *vs* Q3		60/188 (32%)		52/389 (13%)		1.6 *vs* 2.4	0.67 (0.45, 1.01)	6.2 *vs* 5.3
				Q1 *vs* Q4		48/188 (26%)		89/389 (23%)		1.6 *vs* 1.1	1.44 (0.98, 2.11)	6.2 *vs* 33.3
Navarro-Compán 2020[c] [50]	w16	IXE	nr-axSpA	<5 *vs* ≥5	17/41 (42%)	18/55 (32)	7/39 (18%)	13/66 (20)	NR	2.3 *vs* 1.7	1.36 (0.51, 3.62)	4.2 *vs* 8.3
ASAS-PR												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	78/225 (35%)	43/141 (30%)	4/50 (8%)	0/28 (0%)	≤0.05	4.3 *vs* 17.1	0.25 (0.01–4.70)	3.7 *vs* 3.3
				<2 *vs* ≥5–10		42/151 (28%)		5/46 (10.9%)		4.3 *vs* 2.6	1.69 (0.46, 6.16)	3.7 *vs* 5.9
				<2 *vs* >10		66/289 (23%)		9/96 (9.4%)		4.3 *vs* 2.4	1.78 (0.56, 5.68)	3.7 *vs* 7.5
ASDAS-ID												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	61/225 (27%)	30/141 (21%)	3/50 (6%)	0/28 (0%)	NS	4.5 *vs* 11.9	0.38 (0.02, 7.49)	4.7 *vs* 4.7
				<2 *vs* ≥5–10		35/150 (23%)		5/46 (11%)		4.5 *vs* 2.2	2.10 (0.51, 8.71)	4.7 *vs* 8.1
				<2 *vs* >10		62/289 (22%)		9/96 (9%)		4.5 *vs* 2.0	1.97 (0.54, 7.23)	4.7 *vs* 8.3
ASDAS-MI												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	81/225 (36%)	56/141 (40%)	3/50 (6%)	1/28 (4%)	NS	6.0 *vs* 11.1	0.54 (0.06, 5.02)	3.3 *vs* 2.8
				<2 *vs* ≥5–10		64/150 (43%)		3/46 (6%)		6.0 *vs* 6.5	0.92 (0.19, 4.41)	3.3 *vs* 2.8
				<2 *vs* >10		112/289 (39%)		9/96 (9%)		6.0 *vs* 4.1	1.45 (0.40, 5.23)	3.3 *vs* 3.4
BASDAI50												
Rudwaleit 2004 [30]	w12	IFX/ETN	r-axSpA	<10 *vs* ≥10–20	27/37 (73%)	19/33 (58%)	NR	NR	0.003[d]	–	–	–
				<10 *vs* >20		9/29 (31%)		NR		–	–	–
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	137/225 (61%)	89/141 (63%)	10/50 (20%)	1/28 (4%)	NS	3.0 *vs* 17.7	0.17 (0.02, 1.28)	2.4 *vs* 1.7
				<2 *vs* ≥5–10		84/151 (56%)		11/46 (24%)		3.0 *vs* 2.5	1.31 (0.60, 2.85)	2.4 *vs* 3.2
				<2 *vs* >10		159/289 (55%)		15/96 (16%)		3.0 *vs* 3.5	0.86 (0.41, 1.81)	2.4 *vs* 2.5
Navarro-Compán 2020[c] [50]	w16	IXE	nr-axSpA	<5 *vs* ≥5	16/41 (39)	14/55 (26)	3/39 (8)	12/66 (18)	NR	5.1 *vs* 1.4	3.62 (0.95, 13.84)	3.2 *vs* 12.5
SF-36 (PCS) MCID												
Deodhar 2019[c] [48]	w16	SEC	r-axSpA	<2 *vs* ≥2	117/158 (74)	188/269 (70)	56/112 (50)	90/200 (45)	NR	1.5 *vs* 1.6	0.95 (0.73, 1.25)	4.2 *vs* 4.0

(continued)

Table 2. (continued)

Continuous outcomes Study	Time point	Drug	Population	Disease duration (years)	Outcome drug group		Outcome placebo group		*P*-value	Change since baseline		
					Early disease	Established disease	Early disease	Established disease		Early	Established	DiD
Change in ASDAS-CRP												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	–1.7 (–1.9, –1.6)	–1.6 (–1.8, –1.5)	–0.6 (–0.9, –0.4)	–0.1 (–0.4, 0.3)	NS	–1.1	–1.5	0.4
				<2 *vs* ≥5–10		–1.6 (–1.8, –1.5)		–0.4 (–0.7, –0.1)		–1.1	–1.2	0.1
				<2 *vs* >10		–1.6 (–1.7, –1.5)		–0.4 (–0.6, –0.2)		–1.1	–1.2	0.1
Change in BASDAI												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	–32.5 (–35.8, –29.2)	–29.5 (–33.4, –25.6)	–16.2 (–22.1, –10.3)	–4.5 (–12.1, 3.1)	NS	–16.3	–25.0	8.7
				<2 *vs* ≥5–10		–29.6 (–33.4, –25.8)		–13.0 (–19.1, –7.0)		–16.3	–16.6	0.3
				<2 *vs* >10		–29.0 (–31.6, –26.4)		–11.1 (–15.5, –6.7)		–16.3	–17.9	1.6
Change in PGA (0–100)												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	–33.9 (–36.9, –30.9)	–28.9 (–32.4, –25.4)	–17.9 (–23.6, –12.1)	–5.6 (–13.0, 1.8)	0.05[f]	–16.0	–23.3	7.3
				<2 *vs* ≥5–10		–30.8 (–34.2, –27.4)		–16.1 (–22.0, –10.2)		–16.0	–14.7	–1.3
				<2 *vs* >10		–29.1 (–31.4, –26.7)		–13.6 (–18.0, –9.3)		–16.0	–15.5	–0.5
Change in BASFI (0–100)												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	–25.0 (–28.4, –21.6)	–24.2 (–28.1, –20.2)	–12.7 (–18.3, –7.2)	2.3 (–4.8, 9.3)	0.05[e]	–12.3	–21.9	9.6
				<2 *vs* ≥5–10		–23.1 (–26.9, –19.3)		–7.1 (–12.7, –1.4)		–12.3	–16.0	3.7
				<2 *vs* >10		–21.3 (–23.9, –18.6)		–6.6 (–10.8, –2.4)		–12.3	–14.7	2.4
Change in CRP (g/dl)												
Baraliakos 2015 [5]	w12	ETN	r-axSpA	<2 *vs* ≥2–5	–13.0 (–14.7, –11.4)	–13.3 (–15.2, –11.4)	–2.7 (–7.0, –1.5)	1.7 (–3.7, –7.1)	NS	–10.3	–11.6	1.3
				<2 *vs* ≥5–10		–11.7 (–13.5, –9.9)		2.9 (–1.4, –7.3)		–10.3	–8.8	–1.5
				<2 *vs* >10		–12.8 (–14.1, –11.5)		–2.5 (–5.7, –0.6)		–10.3	–10.3	0

Cell colours: green: in favour of early disease; red: in favour of established disease; yellow: not statistically significant. Colour version available online.

[a] Data extracted in a second stage from the FDA website.

[b] AUTHOR PLEASE ADD FOOTNOTE.

[c] Congress abstract.

[d] Interaction *P*-value.

[e] $P \leq 0.05$ for disease duration categories in placebo.

[f] $P \leq 0.05$ for disease duration categories in treatment group.

RR: relative risk; RRR: relative risk ratio; NNTs: number of patients needed to treat; change since baseline, negative values correspond to outcome improvement; DiD: differences in differences, positive values in favour of late disease; ASAS20/40: ASAS response criteria; ASAS-PR: ASAS partial remission; ASDAS-ID: ASDAS Inactive Disease; ASDAS-MI: AS Disease Activity Score Major Improvement criteria; BASDAI50: BASDAI 50 response; SF-36 (PCS) MCID: Short Form-36 (physical component summary), Minimal Clinically Important Difference; ASDAS-CRP: ASDAS-C Reactive Protein; PGA: Patient Global Assessment; BASFI: Bath AS Functional Index; GLM: golimumab 50 mg Q4W; ADA: adalimumab 40 mg Q2W; ETN: etanercept 50 mg/week; SEC: secukinumab 150 mg Q4W; IXE: ixekizumab 80 mg Q4W; IFX: infliximab 5 mg/kg/8 weeks; r-axSpA: radiographic axial SpA; NS: non-significant; NR: not reported.

Table 3. Treatment response in randomized controlled trials comparing early *vs* established disease defined based on radiographic damage in patients with short symptom duration

Study	Time point	Drug	Early *vs* established	Outcome drug group		Outcome placebo group		*P*-value	RR (early *vs* established)	RRR (95% CI)	NNTs (early *vs* established)
				Early	Established	Early	Established				
ASAS20											
Sieper 2016 [42]	w28	IFX	nr-axSpA *vs* r-axSpA	29/40 (72%)	53/61 (87%)	11/16 (69%)	24/33 (73%)	NR	1.0 *vs* 1.2	0.88 (0.56, 1.38)	27.0 *vs* 7.0
Landewé 2020 [43]	w48	CZP	nr-axSpA *vs* r-axSpA	261/329 (79%)	325/407 (80%)	NR	NR	NR	–	–	–
ASAS40											
Song 2013 [41]	w48	ETN	nr-axSpA *vs* r-axSpA	65%	75%	NR	NR	NS	–	–	–
Sieper 2016 [42]	w28	IFN	nr-axSpA *vs* r-axSpA	24/40 (60%)	53/61 (87%)	9/16 (56%)	18/33 (54%)	NR	1.1 *vs* 1.6	0.67 (0.37, 1.22)	27.0 *vs* 3.1
Landewé 2020 [43]	w48	CZP	nr-axSpA *vs* r-axSpA	240/329 (73%)	290/407 (71%)	NR	NR	NR	–	–	–
ASAS-PR											
Song 2013 [41]	w48	ETN	nr-axSpA *vs* r-axSpA	60%	40%	NR	NR	NS	–	–	–
Sieper 2016 [42]	w28	IFX	nr-axSpA *vs* r-axSpA	20/40 (50%)	43/61 (70%)	6/16 (38%)	11/33 (33%)	NR	1.3 *vs* 2.1	0.63 (0.26, 1.50)	8.0 *vs* 2.7
Landewé 2020 [43]	w48	CZP	nr-axSpA *vs* r-axSpA	180/329 (55%)	281/407 (69%)	NR	NR	NR	–	–	–
ASAS-5/6											
Landewé 2020	w48	CZP	nr-axSpA *vs* r-axSpA	195/329 (59%)	227/407 (56%)	NR	NR	NR	–	–	–
ASDAS-ID											
Song 2013 [41]	w12	CZP	nr-axSpA *vs* r-axSpA	40%	40%	NR	NR	NS	–	–	–
Sieper 2016 [42]	w28	IFN	nr-axSpA *vs* r-axSpA	19/36 (53%)	33/57 (58%)	4/13 (31%)	5/30 (17%)	NR	1.7 *vs* 3.4	0.49 (0.15, 1.65)	4.5 *vs* 2.4
Landewé 2020 [43]	w48	CZP	nr-axSpA *vs* r-axSpA	172/327 (53%)	213/407 (52%)	NR	NR	NR	–	–	–
ASDAS-MI											
Song 2013 [41]	w12	CZP	nr-axSpA *vs* r-axSpA	25%	30%	NR	NR	NS	–	–	–
Landewé 2020 [43]	w48	CZP	nr-axSpA *vs* r-axSpA	175 (53%)	238 (58%)	NR	NR	NR	–	–	–
BASDAI50											
Landewé 2020 [43]	w48	CZP	nr-axSpA *vs* r-axSpA	238 (72%)	290 (71%)	NR	NR	NR	–	–	–
BASDAI <3 cm											
Sieper 2016 [42]	w28	IFX	nr-axSpA *vs* r-axSpA	22/36 (61%)	50/58 (86%)	7/13 (54%)	15/30 (50%)	NR	1.1 *vs* 1.7	0.66 (0.33, 1.30)	13.7 *vs* 2.8

Cell colours: green: in favour of early disease; red: in favour of established disease; yellow: not statistically significant. Colour version available online. RR: relative risk; RRR: relative risk ratio; NNTs: number of patients needed to treat; ASAS20/40: ASAS response criteria; ASAS-PR: ASAS partial remission; ASDAS-ID: ASDAS Inactive Disease; ASDAS-MI: AS Disease Activity Score Major Improvement criteria; BASDAI50: BASDAI 50 response; CZP: certolizumab pegol 200 mg Q2W; IFX: infliximab 5 mg/kg Q4W; ETN: etanercept 50 mg/week; nr-axSpA: non radiographic axial SpA; r-axSpA: radiographic axial SpA; NR: not reported; NS: non-significant.

Table 4. Treatment response in RCTs comparing early *vs* established disease defined based on radiographic damage in patients regardless short symptom duration

Study	Time point	Drug	Early *vs* Established	Outcome drug group		Outcome placebo group		*P*-value	RR (early *vs* established)	RRR (95% CI)	NNTs (early *vs* established)
				Early disease	Established disease	Early disease	Established disease				
ASAS20											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	27/46 (59%)	37/65 (57%)	20/50 (40%)	21/57 (37%)	NR	1.5 *vs* 1.6	0.95 (0.53, 1.69)	5.4 *vs* 5.0
ASAS40											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	22/46 (48%)	26/65 (40%)	8/50(16%)	11/57 (19%)	NR	3.0 *vs* 1.4	1.44 (0.57, 3.66)	3.1 *vs* 4.8
ASAS-PR											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	13/46 (28%)	13/65 (20%)	3/50 (6%)	1/56 (2%)	NR	4.4 *vs* 11.4	0.42 (0.04, 4.32)	4.5 *vs* 5.5
ASAS-5/6											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	19/46 (41%)	31/65 (48%)	4/50 (8%)	5/57 (9%)	NR	5.2 *vs* 5.4	0.95 (0.25, 3.59)	3.0 *vs* 2.6
BASDAI50											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	23/46 (50%)	27/65 (42%)	8/49 (16%)	6/57 (10%)	NR	3.1 *vs* 4.0	0.78 (0.27, 2.26)	2.9 *vs* 3.2
Study	**Time point**	**Drug**	**Early *vs* established**	**Outcome drug group**		**Outcome placebo group**		***P*-value**	**Change since baseline**		
				Early disease	Established disease	Early disease	Established disease		Early	Established	DiD
Change in BASDAI											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	−2.8 (2.3)	−2.6 (2.3)	−1.1 (1.8)	−1.1 (1.7)	NR	−1.7	−1.5	−0.2
Change in ASDAS											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	−1.6 (1.2)	−1.8 (1.2)	−0.5 (0.8)	−0.6 (0.8)	NR	−1.1	−1.2	0.1
Change in BASFI											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	−2.1 (2.4)	−1.8 (2.3)	−0.3 (1.9)	−0.8 (1.6)	NR	−1.8	−1.0	−0.8
Change in BASMI											
Landewé 2014 [44]	w12	CZP	nr-axSpA *vs* r-axSpA	−0.5 (0.9)	−0.6 (1.0)	0.1 (0.7)	−0.3 (0.9)	NR	−0.6	−0.3	−0.3

Cells colours: green: in favour of early disease; red: in favour of established disease; yellow: not statistically significant. Colour version available online. RR: relative risk; RRR: relative risk ratio; NNTs: number of patients needed to treat; change since baseline, negative values correspond to outcome improvement; DiD: differences in differences, positive values in favour of late disease; ASAS20/40: ASAS response criteria; ASAS-PR: ASAS partial remission; response; ASDAS: AS Disease Activity Score; CZP: certolizumab pegol 200 mg Q2W; nr-axSpA: non-radiographic SpA; r-axSpA: radiographic axial SpA; NR: not reported.

Consequently, we may not have enough evidence to draw definite conclusions on the treatment response according to symptom duration in axSpA.

The results of the present SLR differ from trials claiming better treatment outcomes in early disease based on a time definition. This discrepancy can be explained by the fact that those studies were simply not designed with the purpose of assessing the interaction between early disease and treatment response, but rather to compare differences between active treatment and placebo in the total population, i.e. regardless of early or established disease. Therefore, such *post hoc* analyses may be subject to biases and may reflect disbalances between the groups in prognostic features as patients were not randomly allocated to the groups at baseline. Additionally, some studies compare the proportion of patients achieving the outcome in early *vs* established axSpA under active treatment, concluding that differences between both groups are attributable to early disease. However, a simple numerical comparison is not enough to answer our main question, not to mention that it also implies overlooking the placebo arm, where the same higher outcome achievement in early disease may take place. Hence, in order to perform a real comparison of the treatment effect of bDMARDs (*vs* placebo) between the two stages of the disease (also referred to as the interaction between treatment and symptom duration) it is elegant and useful to resort to RRR. By dividing the RR of bDMARDs (*vs* placebo) in early axSpA by the RR in established disease, we can truly estimate the relationship between the effect of bDMARDs *vs* placebo in early *vs* established disease, as well as its magnitude, reflected in the value of the RRR. By calculating its 95% CI, we can have an uncertainty around the estimate, allowing assessment of the statistical significance of this between-groups comparison.

As for the comparison between nr-axSpA and r-axSpA, our results are consistent with a meta-analysis conducted by Callhoff *et al.* [54] assessing the difference in BASDAI and BASFI improvement between active treatment and placebo in both groups. After adjustment for year of publication of the RCTs, there were no differences in the effect bDMARDs between the two groups.

Putting the results of this SLR into the context of the ASAS-SPEAR project, a universal definition of early disease to be used for research purposes will allow the comparison of outcomes in early and established disease across studies and therefore the attainment of more solid data to understand the real effect of early treatment. Additionally, well-designed studies will enable determination, based on high-quality evidence, of the best threshold (if any) to define early axSpA, if this is to be defined based on symptom duration.

Despite the lack of robustness of the findings, another key topic is the clarification of our results: lack of differences in treatment effect between early and established disease should not be understood as 'no need to treat patients early'. An adequate treatment should be initiated as soon as patients are diagnosed with axSpA [55]. In fact, the lack of benefit in early treatment would suggest that, unlike other diseases like RA, where a treatment delay leads to worse outcomes with structural and irreversible damage [56–59], a late treatment in axSpA would not straightforwardly indicate that a patient will have a worse prognosis. On the other hand, in order to give a definite answer to this question, we would rather see the impact of early (*vs* late) treatment on structural damage progression. Due to the methodological challenges with the assessment of structural damage progression in axSpA, known to be slow and not optimally captured by the most widely used instruments, it is unlikely that we will be able to set up such a study 'free of bias'. The long-term follow-up that would be needed to answer such a question would pose other methodological challenges like losses to follow-up and confounding, not to mention the unethical issues of performing an RCT with the requested characteristics.

To conclude, this is the first SLR comparing differences in treatment effect between early and established axSpA. Evidence towards better outcomes in early axSpA is very limited, based on three studies, and restricted to nr-axSpA and <5 years symptom duration showing better outcomes compared with established nr-axSpA. The effect of symptom duration on response to treatment needs to be better investigated to allow a definite conclusion. When early axSpA is defined based on disease duration or radiographic damage, no differences in response to treatment are found based upon currently available evidence.

Supplementary data

Supplementary data are available at *Rheumatology* online.

Data availability statement

The data underlying this article will be shared on request to the corresponding author with permission of the ASAS-SPEAR Steering Committee.

Funding

The Assessment of Spondyloarthritis international Society (ASAS) has financially supported the ASAS-SPEAR project.

Disclosure statement: D.C. reports speakers bureau: Bristol Myers Squibb, Pfizer; grant/research support from: Pfizer. D.B. reports speakers bureau: Jannsen, Roche; grant/research support from: Novartis. D.V.d.H. reports consultant of AbbVie, Bayer, BMS, Cyxone, Eisai, Galapagos, Gilead, Glaxo-Smith-Kline, Janssen, Lilly, Novartis, Pfizer, UCB Pharma; employee of: Director of Imaging Rheumatology. R.L. reports consultant of: AbbVie, BMS, Galapagos, Gilead, GSK, Janssen, Lilly, Novartis, Pfizer, UCB. D.P. reports speakers bureau: AbbVie, Bristol-Myers Squibb, Eli Lilly, MSD, Novartis, Pfizer and UCB; consultant of: AbbVie, Biocad, Eli Lilly, Gilead, GlaxoSmithKline, Janssen, MSD, Novartis, Pfizer, Samsung Bioepis and UCB; grant/research support from: AbbVie, Eli Lilly, MSD, Novartis and Pfizer. A.v.T. reports consultant of: Novartis, Galapagos; grant/research support from: Pfizer, UCB, Novartis. L.F.: none declared. V.N.C. reports speakers bureau: AbbVie, Eli Lilly, Janssen, MSD, Novartis, Pfizer, UCB Pharma; consultant of: AbbVie, Eli Lilly, MSD, Novartis, Pfizer, UCB Pharma; grant/research support from: AbbVie and Novartis. S.R. reports speakers bureau: Eli Lilly, MSD, Novartis, UCB; consultant of: AbbVie, Eli Lilly, MSD, Novartis, Pfizer, UCB, Sanofi; grant/research support from: AbbVie, Galapagos, Novartis, Pfizer, UCB.

Acknowledgements

We acknowledge the support of the Assessment of Spondyloarthritis international Society (ASAS) in the development of this project.

D.C. was responsible for study design and set up, data collection, formal analysis, presentation of the data, discussion and interpretation of the findings, and writing of the original draft. D.B. contributed study design, data collection, formal analysis, presentation of the data, discussion and interpretation of the findings, and manuscript revision. D.v.d.H., R.L., D.P. and A.v.T. were involved in review of SLR protocol, methodology, discussion and interpretation of the findings, and manuscript revision. L.F. contributed search strategy development, data collection and manuscript revision. V.N.-C. and S.R. carried out study idea, design and set up, methodology, presentation of the data, discussion and interpretation of the findings, manuscript revision and supervision.

References

1. van Nies JA, Tsonaka R, Gaujoux-Viala C, Fautrel B, van der Helm-van Mil AH. Evaluating relationships between symptom duration and persistence of rheumatoid arthritis: does a window of opportunity exist? Results on the Leiden early arthritis clinic and ESPOIR cohorts. Ann Rheum Dis 2015;74:806–12.
2. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. Arthritis Rheum 2003;48:1771–4.
3. Fragoulis GE, Siebert S. Treatment strategies in axial spondyloarthritis: what, when and how? Rheumatology (Oxford) 2020;59:iv79–89.
4. Robinson PC, Brown MA. The window of opportunity: a relevant concept for axial spondyloarthritis. Arthritis Res Ther 2014;16:109.
5. Baraliakos X, Koenig AS, Jones H *et al.* Predictors of clinical remission under anti-tumor necrosis factor treatment in patients with ankylosing spondylitis: pooled analysis from large randomized clinical trials. J Rheumatol 2015;42:1418–26.
6. Benavent D, Capelusnik D, van der Heijde D *et al.* How is early spondyloarthritis defined in the literature? Results from a systematic review. Semin Arthritis Rheum 2022;55:152032.
7. Luckmann R. Evidence-based medicine: how to practice and teach EBM, 2nd Edition: by David L. Sackett, Sharon E. Straus, W. Scott Richardson, William Rosenberg, and R. Brian Haynes, Churchill Livingstone, 2000. J Intens Care Med 2001;16:155–6.
8. Anderson JJ, Baron G, van der Heijde D, Felson DT, Dougados M. Ankylosing spondylitis assessment group preliminary definition of short-term improvement in ankylosing spondylitis. Arthritis Rheum 2001;44:1876–86.
9. Brandt J, Listing J, Sieper J *et al.* Development and preselection of criteria for short term improvement after anti-TNF alpha treatment in ankylosing spondylitis. Ann Rheum Dis 2004;63:1438–44.
10. Machado P, Landewé R, Lie E *et al.*; Assessment of SpondyloArthritis international Society. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70:47–53.
11. Braun J, Davis J, Dougados M *et al.*; ASAS Working Group. First update of the international ASAS consensus statement for the use of anti-TNF agents in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65:316–20.
12. Lukas C, Landewé R, Sieper J *et al.*; Assessment of SpondyloArthritis international Society. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:18–24.
13. Garrett S, Jenkinson T, Kennedy LG *et al.* A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
14. Calin A, Garrett S, Whitelock H *et al.* A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21:2281–5.
15. Jenkinson TR, Mallorie PA, Whitelock HC *et al.* Defining spinal mobility in ankylosing spondylitis (AS). The Bath AS Metrology Index. J Rheumatol 1994;21:1694–8.
16. Strand V, Singh JA. Improved health-related quality of life with effective disease-modifying antirheumatic drugs: evidence from randomized controlled trials. Am J Manag Care 2008;14:234–54.
17. Kiltz U, van der Heijde D, Boonen A *et al.* Development of a health index in patients with ankylosing spondylitis (ASAS HI): final result of a global initiative based on the ICF guided by ASAS. Ann Rheum Dis 2015;74:830–5.
18. Zhao SS, Pittam B, Harrison NL *et al.* Diagnostic delay in axial spondyloarthritis: a systematic review and meta-analysis. Rheumatology (Oxford) 2021;60:1620–8.
19. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016;5:210.
20. 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.
21. Higgins JP, Altman DG, Gøtzsche PC *et al.*; Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.
22. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010;25:603–5.
23. Hayden JA, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med 2006;144:427–37.
24. Pinson L, Gray GE. Psychopharmacology: number needed to treat: an underused measure of treatment effect. Psychiatr Serv 2003;54:145–6, 154.
25. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003;157:940–3.
26. Rybicki BA, Elston RC. The relationship between the sibling recurrence-risk ratio and genotype relative risk. Am J Hum Genet 2000;66:593–604.
27. de Chaisemartin C, D'HaultfŒuille X. Fuzzy differences-in-differences. Rev Econ Stud 2018;85:999–1028.
28. Sieper J, van der Heijde D, Dougados M *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.
29. Sieper J, Landewé R, Magrey M *et al.* Predictors of remission in patients with non-radiographic axial spondyloarthritis receiving open-label adalimumab in the ABILITY-3 study. RMD Open 2019;5:e000917.
30. 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;63:665–70.
31. Huang F, Gu J, Zhu P *et al.* Efficacy and safety of adalimumab in Chinese adults with active ankylosing spondylitis: results of a randomised, controlled trial. Ann Rheum Dis 2014;73:587–94.
32. Arends S, Brouwer E, van der Veer E *et al.* Baseline predictors of response and discontinuation of tumor necrosis factor-alpha blocking therapy in ankylosing spondylitis: a prospective longitudinal observational cohort study. Arthritis Res Ther 2011;13:R94.
33. Ciurea A, Scherer A, Exer P *et al.*; Rheumatologists of the Swiss Clinical Quality Management Program for Axial Spondyloarthritis. Tumor necrosis factor α inhibition in radiographic and nonradiographic axial spondyloarthritis: results from a large observational cohort. Arthritis Rheum 2013;65:3096–106.
34. Rudwaleit M, Claudepierre P, Wordsworth P *et al.* Effectiveness, safety, and predictors of good clinical response in 1250 patients treated with adalimumab for active ankylosing spondylitis. J Rheumatol 2009;36:801–8.
35. Lord PA, Farragher TM, Lunt M *et al.*; BSR Biologics Register. Predictors of response to anti-TNF therapy in ankylosing spondylitis: results from the British Society for Rheumatology Biologics Register. Rheumatology (Oxford) 2010;49:563–70.

36. Glintborg B, Ostergaard M, Krogh NS *et al.* Predictors of treatment response and drug continuation in 842 patients with ankylosing spondylitis treated with anti-tumour necrosis factor: results from 8 years' surveillance in the Danish nationwide DANBIO registry. Ann Rheum Dis 2010;69:2002–8.
37. Chen MH, Lee MH, Liao HT *et al.* Health-related quality of life outcomes in patients with rheumatoid arthritis and ankylosing spondylitis after tapering biologic treatment. Clin Rheumatol 2018;37:429–38.
38. Yahya F, Gaffney K, Hamilton L *et al.*; for BRITSpA. Tumour necrosis factor inhibitor survival and predictors of response in axial spondyloarthritis-findings from a United Kingdom cohort. Rheumatology (Oxford) 2018;57:619–24.
39. Shimabuco AY, Gonçalves CR, Moraes JCB *et al.* Factors associated with ASDAS remission in a long-term study of ankylosing spondylitis patients under tumor necrosis factor inhibitors. Adv Rheumatol 2018;58:40.
40. Lubrano E, Perrotta FM, Manara M *et al.* Improvement of function and its determinants in a group of axial spondyloarthritis patients treated with TNF inhibitors: a real-life study. Rheumatol Ther 2020;7:301–10.
41. Song IH, Weiß A, Hermann KG *et al.* Similar response rates in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis after 1 year of treatment with etanercept: results from the ESTHER trial. Ann Rheum Dis 2013;72:823–5.
42. Sieper J, Rudwaleit M, Lenaerts J *et al.* Partial remission in ankylosing spondylitis and non-radiographic axial spondyloarthritis in treatment with infliximab plus naproxen or naproxen alone: associations between partial remission and baseline disease characteristics. Rheumatology (Oxford) 2016;55:1946–53.
43. Landewé R, van der Heijde D, Dougados M *et al.* Induction of sustained clinical remission in early axial spondyloarthritis following certolizumab pegol treatment: 48-week outcomes from C-OPTIMISE. Rheumatol Ther 2020;7:581–99.
44. Landewé R, Braun J, Deodhar A *et al.* Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis 2014;73:39–47.
45. Moltó A, Paternotte S, Claudepierre P, Breban M, Dougados M. Effectiveness of tumor necrosis factor α blockers in early axial spondyloarthritis: data from the DESIR cohort. Arthritis Rheumatol 2014;66:1734–44.
46. Corli J, Flipo RM, Philippe P *et al.* Tumor necrosis factor-α inhibition in ankylosing spondylitis and nonradiographic axial spondyloarthritis: treatment response, drug survival, and patient outcome. J Rheumatol 2015;42:2376–82.
47. Kay JG, Deodhar A, Maksymowych W *et al.* Earlier treatment of non-radiographic axial spondyloarthritis with certolizumab pegol results in improved clinical and patient-reported outcomes [abstract]. Arthritis Rheumatol 2019;71.
48. Deodhar A, Boonen A, Ferraccioli G *et al.* FRI0388 Secukinumab improves health-related quality of life in patients with ankylosing spondylitis, irrespective of time since first diagnosis: pooled results from the secukinumab phase 3 trial program. Ann Rheum Dis 2019;78(Suppl 2):878–9.
49. Deodhar AM, Machado P, Meng X, Strand V, Magrey M. Impact of age and disease duration on the response to IL-17A inhibitor (secukinumab) treatment in ankylosing spondylitis: pooled results from the phase 3 MEASURE studies [abstract]. Arthritis Rheumatol 2019;71.
50. Navarro-Compán VM-C, Rahman P, Kronbergs A *et al.* Response to treatment with ixekizumab in patients with active non-radiographic axial spondyloarthritis based on HLA-B27 status and disease duration [abstract]. Arthritis Rheumatol 2020;72.
51. Ørnbjerg LG, Lindström U, Loft A *et al.* Predicting ASDAS inactive disease after 6 months of TNFi treatment in bio-naive axial spondyloarthritis patients treated in clinical practice – results from the EuroSpA collaboration [abstract]. Arthritis Rheumatol 2019;71.
52. FDA. Center for drug evaluation and research-rapid-axSpA 2013 [32–3]. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/125160Orig1s215.pdf.
53. FDA. Center for drug evaluation and research-GO-RAISE. 2009. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2009/125289s000_ClinPharmR_P1.pdf.
54. Callhoff J, Sieper J, Weiß A, Zink A, Listing J. Efficacy of TNFα blockers in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis: a meta-analysis. Ann Rheum Dis 2015;74:1241–8.
55. van der Heijde D, Ramiro S, Landewé R *et al.* 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017;76:978–91.
56. van Everdingen AA, Jacobs JW, Siewertsz Van Reesema DR, Bijlsma JW. Low-dose prednisone therapy for patients with early active rheumatoid arthritis: clinical efficacy, disease-modifying properties, and side effects: a randomized, double-blind, placebo-controlled clinical trial. Ann Intern Med 2002;136:1–12.
57. Borg G, Allander E, Lund B *et al.* Auranofin improves outcome in early rheumatoid arthritis. Results from a 2-year, double blind placebo controlled study. J Rheumatol 1988;15:1747–54.
58. Egsmose C, Lund B, Borg G *et al.* Patients with rheumatoid arthritis benefit from early 2nd line therapy: 5 year followup of a prospective double blind placebo controlled study. J Rheumatol 1995;22:2208–13.
59. van Nies JA, Krabben A, Schoones JW *et al.* What is the evidence for the presence of a therapeutic window of opportunity in rheumatoid arthritis? A systematic literature review. Ann Rheum Dis 2014;73:861–70.