EXTENDED REPORT

Preliminary definitions of 'flare' in axial spondyloarthritis, based on pain, BASDAI and ASDAS-CRP: an ASAS initiative

Laure Gossec,[1,2] Agnès Portier,[1,3,4] Robert Landewé,[5] Adrien Etcheto,[3,4] Victoria Navarro-Compán,[6,7] Féline Kroon,[6] Désirée van der Heijde,[6] Maxime Dougados[3,4]

Handling editor Tore K Kvien

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2015-208593).

For numbered affiliations see end of article.

Correspondence to
Professor Laure Gossec, Hôpital Pitié-Salpétrière, Service de Rhumatologie, 47-83, boulevard de l'Hôpital, Paris 75013, France; laure.gossec@aphp.fr

Received 14 September 2015
Revised 4 December 2015
Accepted 16 January 2016
Published Online First 4 February 2016

ABSTRACT

Introduction Flares may be used as outcomes in axial spondyloarthritis (axSpA) trials or observational studies. The objective was to develop a definition for 'flare' (or worsening) in axSpA, based on validated composite indices, to be used in the context of clinical trial design.

Methods (1) Systematic literature review of definitions of 'flare' in published randomised controlled trials in axSpA. (2) Vignette exercise: 140 scenarios were constructed for a typical patient with axSpA seen at two consecutive visits. Each scenario included a change in one of the following outcomes: pain, Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), BASDAI plus C-reactive protein (CRP) or Ankylosing Spondylitis Disease Activity Score (ASDAS)-CRP. Each Assessment of Spondyloarthritis (ASAS) expert determined if every scenario from a random sample of 46 scenarios was considered a flare (yes/no). Receiver-operating characteristic (ROC) analyses were applied to derive optimal cut-off values. (3) ASAS consensus was reached.

Results (1) The literature review yielded 38 studies using some definition of 'flare', with 27 different definitions indicating important heterogeneity. The most frequent definitions were based on BASDAI changes or pain changes. (2) 121 ASAS experts completed 4999 flare assessments. The areas under the ROC curves were high (range: 0.88–0.89). Preliminary cut-offs for pain (N=3), BASDAI (N=5) and ASDAS-CRP (N=4) were chosen, with a range of sensitivity 0.60–0.99 and range of specificity 0.40–0.94 against the expert's opinions.

Conclusions This data-driven ASAS consensus process has led to 12 preliminary draft definitions of 'flare' in axSpA, based on widely used indices. These preliminary definitions will need validation in real patient data.

INTRODUCTION

The natural course of axial spondyloarthritis (axSpA) includes periods of flares and remission.[1] Flares are an important attribute of disease activity, and assessment of flares is useful in clinical practice and in clinical trials to better understand disease status and treatment efficacy. In the context of clinical trials, the assessment of flares is necessary in two situations: in 'flare-design trials', trial treatment is introduced only in case of flare being the consequence of interruption of the ongoing/previous treatment (eg, in axSpA if non-steroidal anti-inflammatory drugs (NSAIDs) have been stopped)[2]; and in tapering or discontinuation trials, if the treatment (eg, tumour necrosis factor inhibitors (TNFis)) is (usually progressively) tapered or discontinued in patients being in a stable disease activity state, and the outcome measure is (time to) flare.[3 4]

Thus the concept of flare—or disease activity worsening—needs to be well established in axSpA. This is particularly important since one can anticipate an increasing number of studies will concern drug discontinuation in patients being in remission or low disease activity on treatment. Criteria to define 'flare' may help harmonising trial designs for new clinical trials and may lead to better assessment of axSpA and its fluctuations. However, to date, a broadly accepted definition of 'flare' in axSpA is lacking. Indeed, a succinct check of flare definitions used in published trials indicates important heterogeneity.

The Assessment of Spondyloarthritis (ASAS) group is an international, independent group of experts of spondyloarthritis (SpA) with a methodological focus, which has developed and validated most of the criteria and outcome measures currently used in SpA clinical trials.[5–7] The ASAS group has decided to explore the definition of 'flare' in axSpA. Ongoing work on flares in rheumatoid arthritis (RA) is exploring differences in the perception of flares by physicians and patients, with the objective to develop a specific outcome measure, that is, a new questionnaire, to assess flares in RA.[8 9] There are previously published studies on the perception of flare by the patient in SpA.[10–12] However, in the present project, it was decided not to explore the patients' perspective per se, but rather to focus on the definition of 'flare' based on validated outcomes already widely used to assess disease activity in axSpA, as has recently been done in a French study.[13]

The aim of this project was to develop a consensus definition of 'flare' (or worsening) in axSpA, based on validated composite indices, to be used in clinical trial designs and designs of longitudinal studies.

MATERIAL AND METHODS

This project had two main steps to collect data: a systematic literature review (SLR) and a case-vignette exercise. This was followed by a consensus step.

▶ http://dx.doi.org/10.1136/annrheumdis-2016-209218

To cite: Gossec L, Portier A, Landewé R, *et al*. *Ann Rheum Dis* 2016;**75**: 991–996.

Systematic literature review

Data retrieval

First, to gain an overview of flares, studies specifically focusing on flares in patients with axSpA, with any or no intervention, were searched for in Medline Pubmed and Embase in May 2014. The key words were derived from 'ankylosing spondylitis' and 'flare, exacerbation, relapse, recurrence, clinical reactivation'.

A second SLR was performed to collect all the definitions of 'flare' used in randomised controlled trials (RCTs) of NSAIDs or TNFi in patients with axSpA, up to May 2014. The search was based on two previous systematic reviews and updated in Medline PubMed, Embase and Cochrane for articles published in English, German, French or Spanish. Unpublished RCTs from main rheumatology congress abstracts for 2012–2014 and ongoing trials from the website http://www.clinicaltrials.gov were also analysed. The key words used were derived from 'ankylosing spondylitis' and 'clinical trials'. The search strategy and the full key words are shown in online supplementary table S1.

Data selection

One investigator (AP) selected all the studies referring to the concept of flare in adult patients with axSpA.

Data extraction

General data regarding study characteristics and specific flare data were collected. The outcome of interest was the definition used for 'flare'. If present, information was collected about the instrument used, the cut-off level if flare was measured by a combination of several instruments or as a single instrument only and if flare was conceptualised as a relative change, an absolute change or an absolute value (status).

Analysis was descriptive and included the instrument used to define 'flare', use of one instrument or of a combination, cut-off used to determine flare, use of a relative or absolute change or use of an absolute value.

Vignette exercise

To assess ASAS members' opinions on what constitutes a flare in axSpA, a case-vignette exercise was conducted. Vignettes are brief written case histories of a fictitious patient based on a realistic clinical situation accompanied by one or more questions that explore what a physician would think if presented with the actual patient.[14]

Development of the case-vignettes

The case-vignettes were designed by three authors (LG, AP and MD) based on only one scenario. Full information is given in online supplementary table S3. It was decided to use the case of a 32-year-old man with a well-established diagnosis of axSpA in order to avoid diagnostic discussions. In the scenario, the patient had visits at two successive time points, and a description of the patient's status at both time points was given using results of scores. It was decided that flare would be defined as a change in status between the two time points, that is, a flare is an absolute change between two values: the observed value of the outcome at the time of the flare minus the referral value (previous status before the flare). The scores used here were: (a) patient-reported pain numerical rating scores (pain due to axSpA, range 0–10); (b) Bath Ankylosing Spondylitis Disease Activity Index (BASDAI[15] range 0–10); (c) C-reactive protein (CRP) as a continuous result (in milligram per litre), coupled with change in BASDAI; and (d) the Ankylosing Spondylitis Disease Activity Score—CRP[16] (ASDAS-CRP) as a global score. For illustrative purposes, the elements of the ASDAS-CRP were shown for each ASDAS result: the ASDAS includes back pain, duration of morning stiffness, patient global assessment, peripheral pain/swelling and CRP.[16–18]

The patient's initial status (referral value of the outcome) varied from no symptoms to moderate/high disease activity (eg, pain level of 6/10), thus excluding very high initial values, since it was considered that definitions of 'flares' are only relevant for patients initially not in high/very high disease activity. Many possible steps of worsening in the patient's disease activity status were constructed; in the end, 140 vignettes were designed (see table 1 and online supplementary table S3). An example of a vignette for BASDAI is the following: 'A 32-year-old man with a well-established diagnosis of axSpA consults you at two successive time points. In comparison with the previous visit and according to the following data, and all other things being equal (physical examination, CRP and NSAID intake), do you consider this patient is flaring at the second visit? Yes or No. Please give an answer (yes or no) even if you are unsure'.

Initial (first visit) BASDAI (0–10): 2; final (second visit) BASDAI (0–10): 4; Flare: Yes/No.

Initially, variations in CRP alone, as well as in NSAID intake (ie, 65 additional vignettes), were also constructed but were not retained for the final definitions since the group considered that isolated variations in acute phase reactants or in NSAID intake, without changes in any other parameters, were unlikely to reflect a flare. These results are therefore not presented here.

The timeframe between the two visits was not determined to allow better external validity of the definition.

Distribution of the vignettes

All the 159 ASAS experts were asked to assess a sample of 46 vignettes between July and December 2014; each sample was intentionally constructed to include vignettes for each outcome and a distribution of changes in status. The ASAS experts were asked to answer for each vignette if the patient was considered flaring (yes/no).

Table 1 The outcome changes used in the vignette exercise

Outcomes	BASDAI (0–10)	Pain due to axSpA (numeric rating scale 0–10)	ASDAS score (range, 0.6->5)	BASDAI and CRP combination
Initial level of the outcome at the first visit of the patient	0–6	0–6	0.6–2.0	BASDAI of 2 and CRP of <6 mg/L, 8 mg/L or 20 mg/L
Possible worsenings at the second visit of the patient	Increases in steps of 1 point	Increases in steps of 1 point	Increases in steps 0, 0.3, 0.6, 0.9, 1.1, 1.5	BASDAI of 4 or 5 and CRP increases of 5–20 mg/L
Total number of vignettes: 140	49	49	24	18

ASDAS, Ankylosing Spondylitis Disease Activity Score; axSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C-reactive protein.

Analysis

For each outcome separately, the vignettes were analysed per stratum of change in outcome, that is, for an absolute change of outcome of at least X (thus all vignettes with a BASDAI increase of at least three points were analysed together, then all vignettes with an increase of at least four points and so on). The absolute change in each outcome was then coupled to the value of the variable at the time before the flare (referral value) and the value observed at the time of flare (eg, change in pain of at least 2 points and pain value at time of flare of at least 4 points on a 0–10 scale).

Using the outcome values as the test, and the 'flare-judgement' by the rheumatologist as the 'gold-standard', sensitivity and specificity could be calculated for each of the outcomes and receiver-operating characteristic (ROC) curves were constructed. Areas under the ROC curve were calculated and optimal cut-off values for defining a 'flare' were established. The corresponding sensitivities, specificities, positive predictive values (PPV) and negative predictive values (NPV) were then calculated. For example, the sensitivity is the proportion with a BASDAI change ≥X calculated among those considered in flare by the physician. The specificity is the proportion with a BASDAI change <X calculated among those considered not in flare by the physician. The PPV is the proportion with a flare calculated among those who have a BASDAI change ≥X, and the NPV is the proportion with no flare among those who have a BASDAI change <X.

Final consensus

Results were presented to the ASAS experts during a plenary workshop in January 2015 and consensus on a preliminary set of draft definitions was reached.

RESULTS

SLR of definitions used for flare in axSpA studies

A total of 1013 articles initially screened resulted in 38 studies using some definition of 'flare' in axSpA (see online supplementary table S2). There were 23 RCTs proposing definitions of 'flares', assessing either NSAIDs (N=16) or TNFi (N=7): 19 of them concerned flares between screening and baseline, and 4 concerned flares after drug discontinuation. Of these RCTs, 11 (65%) were published over the last 2 years or were ongoing studies found in clinicaltrials.gov. Additionally, there were 15 studies referring specifically to flares: 8 were trials, 3 were qualitative studies and 4 had another study design.

The 38 studies used 27 different definitions of 'flare' (table 2). The frequency of flares using these definitions was not always reported but when reported, ranged from 7% to 91% (see online supplementary table S2). The two most frequent definitions used were: absolute BASDAI ≥4/10 with absolute physician assessment ≥4/10 used in six studies, and increase in pain ≥30% with absolute pain ≥4/10 used in six studies.

Overall, all 38 (100%) studies with 'flare' definitions used patient-reported outcomes of which 17 (45%) used BASDAI (table 2). BASDAI was used to define flares, either alone (N=7, 41% of 17 studies), or in combination with other instruments (N=10, 59% of 17 studies). Of note, in the literature a flare defined by BASDAI was generally based on a change of at least 1 or 2 points on a 0–10 scale.

Pain was used in 14 (37%) articles to define 'flares', either alone (N=10), or in combination with other instruments (N=4).

ASDAS was used only once to define 'flare' using a cut-off of 2.1 (absolute value).

Five studies (13%) used elements of physical assessment and four (10%) used acute phase reactants to define 'flares' (table 2).

Vignette exercise and final consensus

Of the 159 ASAS members, 121 (76%) completed the exercise (some of them partly), yielding a total of 4999 responses to analyse. The analyses and the consensus process led to 12 preliminary definitions of flare; the performances of these different definitions are shown in table 3 and ROC curves are presented as online supplementary figure S1. Further information is given below.

Pain

The prevalence of the event 'flare' was 63.1% (387 of 613 answers) in the pain vignettes. The ROC curve allowed the selection of two cut-offs for pain variations (on a 0–10 pain scale), with best sensitivity/specificity trade-offs: increase in pain ≥2 points and increase in pain ≥3 points. For these two cut-offs, performances were calculated for different referral (first visit) pain values and observed (second visit) pain values.

The resulting figures (not shown) indicated (a) considering a pain change ≥2 points, more than 70% of the doctors will consider there is a flare if the referral level of pain is ≤4. (b) Considering a pain change ≥2 points, more than 60% of the doctors will consider there is a flare if the final value is ≥4. (c) Considering a pain change ≥3 points, more than 80% of the doctors will consider there is a flare if the referral level of pain is £4. (d) Considering a pain change ≥3, more than 80% of the doctors will consider there is a flare if the final value is ≥5.

Based on these results, and as the referral value defines the context of the study whereas the observed value at the time of the flare defines the flare, it was proposed to keep two preliminary definitions based on pain: (a) an increase in pain of ≥2 *and* an observed value at the time of the flare of ≥4; (b) an increase in pain of ≥3. The performances of these cut-off values are given in table 3. Additional discussions during the consensus process led us to propose the following combined definition: if the observed value is ≥4, a 'flare' is defined as an increase in pain ≥2 points, otherwise, flare is defined as an increase in pain ≥3 points (table 3).

BASDAI

The prevalence of the event 'flare' was 68.1% (421 of 618 answers) in the BASDAI vignettes. The ROC curve allowed the selection of two cut-offs for BASDAI (on a 0–10 scale): increase in BASDAI ≥2 points and increase in BASDAI ≥3 points. For these two cut-offs, the performances were again calculated for different referral and observed values. (a) Considering a BASDAI change ≥2, more than 80% of the doctors will consider there is a flare if the referral BASDAI is ≤4. (b) Considering a BASDAI change ≥2, more than 60% (or 70%) of the doctors will consider there is a flare if the observed value is ≥4 (or 5). (c) Considering a BASDAI change ≥3, more than 80% of the doctors will consider there is a flare if the referral BASDAI is ≤4. (d) Considering a BASDAI change ≥3, more than 70% of the doctors will consider there is a flare if the observed value is ≥4 or 5.

Thus the selected preliminary cut-offs for BASDAI are based on an increase of at least two or at least three points, with or without an observed value of at least four (table 3). An additional (combined) definition was derived during the consensus process as follows: if the observed value of BASDAI is ≥4,

Table 2 The 27 definitions of flares for axSpA found in 38 articles

Type of outcome	Outcome	Number of articles (% of 38 studies)	Outcome used alone or in combination, to define flares	Cut-off used (N articles concerned)
Composite indices	BASDAI (/10)	17 (45)	Combination N=10 Alone N=7	Abs. value ≥4 (N=9) Abs. value ≥3 (N=1) Abs. change 1/10 (N=2) Rel. change 80% or abs. change 2/10 (N=2) Abs. change 1.5/10 (N=1) Rel. change 60% (N=1)
	ASDAS	1 (2.6)	Alone	Abs. value ≥2.1 (N=1)
	ASAS 40 response	1 (2.6)	Alone	Loss (N=1)
Isolated patient-reported outcome	Pain (0–100 mm)	14 (37)	Alone N=10 Combination N=4	Abs. value ≥40 mm and increase of 30% (N=7) Abs. value ≥40 mm and increase of 30% (N=2) Rel. change 50% (N=1) Abs. change 2/10 (N=1) No cut-off (N=3)
	Morning stiffness	5 (13.2)	Combination N=5	Presence (no cut-off) (N=4) Abs. value ≥30 min (N=1)
	Patient global assessment (0–10)	2 (5.2)	Combination N=2	Abs. value ≥4/10 (N=1) Rel. change ≥2/10 (N=1)
	NSAID intake	1 (2.6)	Alone	Presence
Physician assessment or laboratory value	Physician global assessment (0–10)	7(18.4)	Combination N=7	Absolute value ≥4 (N=7)
	Physical assessment	5 (13.2)	Combination N=5	Restriction (N=4) Abs. change: decreased Schöber index (≥1 cm), decreased chest expansion (≥1 cm), increased fingertips to floor distance (≥5 cm)
	Acute phase reactants	4 (10.4)	Combination N=4	Presence (N=3) ESR ≥28 mm (N=1)
	Peripheral or extra-articular manifestation	2 (5.2)	Combination N=2	Presence (N=2)

Abs: absolute; ASAS, Assessment of Spondyloarthritis; ASDAS, Ankylosing Spondylitis Disease Activity Score[16]; axSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index[14]; ESR, erythrocyte sedimentation rate; NSAIDs, non-steroidal anti-inflammatory drugs; Rel: relative.

'flare' is defined as an increase in BASDAI ≥2 points; otherwise, 'flare' is defined as an increase in BASDAI ≥3 points (table 3).

BASDAI+CRP

In the BASDAI+CRP vignettes overall, the prevalence of 'flare' was 77.6% (662 of 852 answers). Not unexpectedly, the analyses suggested a greater role of CRP in defining a flare when the change in BASDAI was ≥2 points than when the change in BASDAI was ≥3 points. In addition, in patients in whom there was no increase of CRP more flares were defined by the physician if the referral value of CRP was abnormal (data not shown). The final decisions that were made were to not propose

Table 3 The 12 ASAS-selected preliminary draft definitions of flare with their performances in the vignette exercise

Instrument	AUC	Flare definition	Se	Spe	PPV	NPV
Pain (0–10)	0.86	Δ pain ≥2 AND final value ≥4	0.99	0.30	0.76	0.97
		Δ pain ≥3	0.95	0.69	0.83	0.88
		If observed value is ≥4: Δ pain ≥2 points, otherwise: Δ pain ≥3 points	0.97	0.56	0.79	0.92
BASDAI (0–10)	0.86	Δ BASDAI ≥2 points	0.99	0.40	0.78	0.94
		Δ BASDAI ≥2 points AND final value ≥4	0.99	0.32	0.81	0.92
		Δ BASDAI ≥3 points	0.92	0.70	0.87	0.80
		Δ BASDAI ≥3 points AND final value ≥4	0.94	0.63	0.88	0.79
		If observed value is ≥4, Δ BASDAI ≥2 points, otherwise: Δ BASDAI ≥3 points	0.94	0.54	0.81	0.80
ASDAS-CRP	0.89	Δ ASDAS ≥0.6	0.97	0.65	0.75	0.96
		Δ ASDAS ≥0.9	0.85	0.87	0.87	0.85
		Δ ASDAS ≥1.1	0.60	0.94	0.93	0.69
		Δ ASDAS ≥0.6 AND observed ASDAS ≥1.3	0.97	0.59	0.78	0.93

ASAS, Assessment of Spondyloarthritis; ASDAS, Ankylosing Spondylitis Disease Activity Score; AUC: area under the receiver-operating characteristic curve; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; CRP, C-reactive protein; PPV and NPV, positive and negative predictive values; Se: sensitivity; Spe: specificity; Δ: change.

the association of a change in BASDAI and a change in CRP as a preliminary definition for flare, but rather to focus on the ASDAS that aggregates this information into one score.

ASDAS-CRP

The prevalence of the event 'flare' was 51.4% (591 of 1150 answers in the ASDAS-CRP vignettes). The ROC curve allowed the selection of three cut-offs for ASDAS-CRP changes: increase in ASDAS-CRP ≥0.6, 0.9 or 1.1. For these three cut-offs, the performances were calculated for different referral and observed values. (a) In contrast to pain and BASDAI, there was no effect of the referral value on the performance of the changes in ASDAS-CRP to define a 'flare'. (b) Regarding the observed values of ASDAS-CRP at the time of flare, there was also no clear effect of this observed ASDAS value on the performance of the cut-offs. Of note, however, only a few vignettes addressed this issue. Based on expert opinion only, an additional preliminary definition of 'flare' based on change in ASDAS associated with an observed value (at the time of flare) of ≥1.3 (ie, not being in inactive disease[18]) was added (table 3).

DISCUSSION

This consensus process, instigated by the ASAS group, has led to 12 preliminary definitions of 'flare' in axSpA, based on widely used indices. Further steps will allow the assessment of these preliminary definitions on real patient data in order to select the most relevant definition(s). This work is important in the context of clinical trial design, for example, for designing tapering trials, to better define 'flares' in future clinical studies.

The initial objective of this initiative was to define a single definition for 'flare' in axSpA. However, a discrepancy was found between the definitions of 'flare' used in the literature and the results of the 'case-vignettes' (in particular, the thresholds to define a 'flare' in the 'case-vignettes' were higher than the thresholds found in the literature). This led ASAS to decide that it was too early to propose a single definition of 'flare'. However, based on the results of both the systematic literature research and the vignette exercise, we are able to focus future studies on 12 potential definitions of 'flare'.

The strengths of this study include an extensive literature review, an extensive vignette process and a strong consensus process, within a well-recognised group of experts in axSpA. A weakness of this study is the limitation of the scenario which does not allow discussions of flares in different subgroups (eg, men vs women; or patients with extra-articular manifestations vs those without). However, the objective of this study was to obtain one simple and uniform definition for 'flare' to be used mainly in clinical trials and studies rather than multiple definitions to be applied in different contexts. Vignette exercises have limitations too, since they only reflect a part of all potential information collected in a real patient/physician consultation; in this case, the vignettes were by nature artificial since patients were considered to show variation in only one outcome, all other things being equal, which is not usually the case in clinical practice. However, vignette exercises are well-recognised ways of obtaining input from many participants.[19 20]

The outcomes chosen in the present initiative can be discussed. BASDAI and pain were selected because these were the two most frequently used instruments in the literature to define 'flares' in axSpA. The ASDAS score was selected because this is a recent instrument validated in axSpA.[16 18] As the ASDAS-CRP is the instrument of choice proposed by ASAS, only ASDAS-CRP (not ASDAS based on the erythrocyte sedimentation rate) was used. CRP was selected because a number of studies used this instrument to assess flares in axSpA. However the interpretation of CRP variations alone (ie, in the absence of concomitant changes in symptoms) was difficult, giving rise to discussions, for example, in case of concomitant infections. Finally, NSAID intake was initially explored to be used in a 'flare' definition, since it may reflect a worsening of the disease, but the interpretation of isolated changes in NSAID intake was very complex.[21] In this vignette exercise, initial levels of symptoms were low to moderate/high since pain could, for example, start at 6/10. In clinical studies, however, most patients will start at low levels, for example, remission. This study does not explore the patient's perspective on flares. Ongoing work in RA has shown that patients and physicians have different perspectives on flares in that disease.[9 22] In axSpA also, it appears patients and physicians may value disease activity differently.[10–12] However, the objective here was not to develop a new score focusing on flares, but rather to define an optimal cut-off value corresponding to a flare or a disease worsening, and applicable to widely used and well-validated outcome measures reflecting disease status in axSpA. It is arguable if a 'flare' can be defined solely as a worsening of disease activity. In the present study we assumed a 'flare' would indeed be best defined as disease worsening. Of note, we did not give any indication, in the vignette exercise, to the ASAS experts of what they should consider to be a flare (eg, worsening necessitating a treatment change), which may have increased the variability in our results.

For the outcomes used in the present study, cut-off values to define improvement have already been defined.[23] However, it is known that minimal clinically important differences are not of the same magnitude when defining an improvement and a worsening. In this regard, this innovative initiative is very much in keeping with the ASAS objectives that aim to provide data-driven approaches to SpA measurement and measure interpretation.

This study focused on the definition of a clinically relevant change in a specific outcome measure reflecting a worsening/deterioration/flare of the disease (ie, minimal clinically important deterioration, MCID), keeping in mind that previously reported studies have proposed definitions of a clinically relevant change reflecting an improvement of the disease (ie, minimal clinically important improvement, MCII). It has been shown in different diseases and for different outcome measures that, for a specific outcome measure of a specific disease, the MCID is usually lower than the MCII.[24]

For example in RA, a change of at least 1.2 in the Disease Activity Score DAS28-ESR is usually considered an MCII and a change of at least 0.6 an MCID.[25] In the field of axSpA, an absolute change in BASDAI of at least 2 points or a relative change of at least 50% have been proposed as an MCII.[26]

Concerning ASDAS-CRP, changes of at least 1.1 and 2.0 have been proposed to define a clinically important improvement (which is in the current context similar to the MCII) and a major improvement, respectively.[18] If we accept the concept that for a specific outcome measure the MCID is at a lower level than the MCII, in our study, the data provided by the SLR might be more relevant than the data from the case-vignette study. The discrepancies observed in our study between the SLR and the case-vignette study might be explained by the fact that the participants in the study (all experts in SpA) were aware of the proposed MCII and unconsciously applied these cut-offs when evaluating a specific scenario.

When discussing flares, the referral status (ie, the patient's status at the time before flare) was arbitrarily defined as a favourable (low activity) status. Indeed, it does not seem rational

to define 'flares' for patients who are already in high disease activity. The referral status can be inactive disease, remission or PASS (Patient Acceptable Symptom State).[18 27] The present study does not define the referral status precisely, in order to allow for better generalisability.

The durability of the status of flare was not explored in the present vignette exercise, but ASAS members felt that a 'flare necessitating treatment intensification' might be defined as a flare observed at least 2 weeks apart or at least at two consecutive visits. This remains to be further explored.

In conclusion, the preliminary definitions of 'flare' given in the present work will now need to be validated on real patient data.

Author affiliations
[1]Sorbonne Universités, UPMC Univ Paris 06, Institut Pierre Louis d'Epidémiologie et de Santé Publique, GRC-UPMC 08 (EEMOIS), Paris, France
[2]Department of Rheumatology, AP-HP, Pitié Salpêtrière Hospital, Paris, France
[3]Medicine Faculty, Paris Descartes University, Paris, France
[4]Rheumatology B Department, APHP, Cochin Hospital, Paris, France
[5]Department of Clinical Immunology & Rheumatology, Amsterdam Rheumatology Center, Amsterdam and Atrium Medical Center, Heerlen, The Netherlands
[6]Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands
[7]Rheumatology Department, University Hospital La Paz, Madrid, Spain

Contributors MD, LG, DH and RL designed the study and oversaw the vignette exercise. AP performed the systematic literature review with the help of FK and VN-C. AP, LG and AE performed the vignette exercise. AE performed the statistical analyses for the vignette exercise. All authors participated in drafting the article and gave final approval.

Funding Supported by ASAS.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The full results of the vignette exercise are available from the last author MD.

REFERENCES

1 Mengshoel AM. Living with a fluctuating illness of ankylosing spondylitis: a qualitative study. *Arthritis Rheum* 2008;59:1439–44.
2 Trijau S, Avouac J, Escalas C, *et al*. Influence of flare design on symptomatic efficacy of non-steroidal anti-inflammatory drugs in osteoarthritis: a meta-analysis of randomized placebo-controlled trials. *Osteoarthritis Cartilage* 2010;18:1012–18.
3 Sieper J, Lenaerts J, Wollenhaupt J, *et al*. Maintenance of biologic-free remission with naproxen or no treatment in patients with early, active axial spondyloarthritis: results from a 6-month, randomised, open-label follow-up study, INFAST Part 2. *Ann Rheum Dis* 2014;73:108–13.
4 Cantini F, Niccoli L, Cassarà E, *et al*. Sustained maintenance of clinical remission after adalimumab dose reduction in patients with early psoriatic arthritis: a long-term follow-up study. *Biol Targets Ther* 2012;6:201–6.
5 Bautista-Molano W, Navarro-Compán V, Landewé RBM, *et al*. How well are the ASAS/OMERACT Core Outcome Sets for Ankylosing Spondylitis implemented in randomized clinical trials? A systematic literature review. *Clin Rheumatol* 2014;33:1313–22.
6 Sieper J, Rudwaleit M, Baraliakos X, *et al*. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 2009;68(Suppl 2):ii1–44.
7 Hermann KGA, Baraliakos X, van der Heijde DMFM, *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.
8 Bartlett SJ, Bykerk VP, Cooksey R, *et al*. Feasibility and domain validation of rheumatoid arthritis (RA) flare core domain set: report of the OMERACT 2014 RA flare group plenary. *J Rheumatol* 2015;42:2185–9.
9 Bartlett SJ, Hewlett S, Bingham CO, *et al*. Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. *Ann Rheum Dis* 2012;71:1855–60.
10 Cooksey R, Brophy S, Gravenor MB, *et al*. Frequency and characteristics of disease flares in ankylosing spondylitis. *Rheumatol Oxf Engl* 2010;49:929–32.
11 Brophy S, Calin A. Definition of disease flare in ankylosing spondylitis: the patients' perspective. *J Rheumatol* 2002;29:954–8.
12 Stone MA, Pomeroy E, Keat A, *et al*. Assessment of the impact of flares in ankylosing spondylitis disease activity using the Flare Illustration. *Rheumatol Oxf Engl* 2008;47:1213–18.
13 Godfrin-Valnet M, Puyraveau M, Prati C, *et al*. Flare in spondyloarthritis: Thresholds of disease activity variations. *Jt Bone Spine Rev Rhum* 2015;82:192–5.
14 Veloski J, Tai S, Evans AS, *et al*. Clinical vignette-based surveys: a tool for assessing physician practice variation. *Am J Med Qual* 2005;20:151–7.
15 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.
16 Lukas C, Landewé R, Sieper J, *et al*. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis* 2009;68:18–24.
17 Machado P, Navarro-Compán V, Landewé R, *et al*. Calculating the ankylosing spondylitis disease activity score if the conventional c-reactive protein level is below the limit of detection or if high-sensitivity c-reactive protein is used: an analysis in the DESIR cohort. *Arthritis Rheumatol* 2015;67:408–13.
18 Machado P, Landewé R, Lie E, *et al*. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. *Ann Rheum Dis* 2011;70:47–53.
19 Banks J, Hollinghurst S, Bigwood L, *et al*. Preferences for cancer investigation: a vignette-based study of primary-care attendees. *Lancet Oncol* 2014;15:232–40.
20 Das J, Kwan A, Daniels B, *et al*. Use of standardised patients to assess quality of tuberculosis care: a pilot, cross-sectional study. *Lancet Infect Dis* 2015;15:1305–13.
21 Dougados M, Simon P, Braun J, *et al*. ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. *Ann Rheum Dis* 2011;70:249–51.
22 Hewlett S, Sanderson T, May J, *et al*. 'I'm hurting, I want to kill myself': rheumatoid arthritis flare is more than a high joint count—an international patient perspective on flare where medical help is sought. *Rheumatol* 2012;51:69–76.
23 Pavy S, Brophy S, Calin A. Establishment of the minimum clinically important difference for the bath ankylosing spondylitis indices: a prospective study. *J Rheumatol* 2005;32:80–5.
24 Tubach F, Wells GA, Ravaud P, *et al*. Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues. *J Rheumatol* 2005;32:2025–9.
25 Curtis JR, Yang S, Chen L, *et al*. Determining the minimally important difference in the clinical disease activity index for improvement and worsening in early rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2015;67:1345–53.
26 van der Heijde D, Sieper J, Maksymowych WP, *et al*. 2010 Update of the international ASAS recommendations for the use of anti-TNF agents in patients with axial spondyloarthritis. *Ann Rheum Dis* 2011;70:905–8.
27 Wariaghli G, Allali F, Idrissi Z, *et al*. Evaluation and stability of the Patient Acceptable Symptom State (PASS) over time in patients with ankylosing spondylitis. *Clin Exp Rheumatol* 2012;30:106–9.

Online supplementary Table 1: the search strategy for the systematic literature review (SLR) to identify flare definitions used in axSpA studies

Overview	The SLR was conducted in 2 phases. The 1st phase was focused on all randomized controlled trials (RCTs) of NSAIDs or antiTNF in axial SpA patients. The 2nd phase consisted on a research of specifically flares-centered studies in axial SpA.
Data Sources and Searches step 1: RCTs	All the RCTs and quasi RCTs using NSAIDs or antiTNF in axial SpA patients were reviewed. First, 2 recent SLRs concerning NSAIDs in axial SpA (REF 1) and concerning antiTNF in axial SpA (REF 2) whatever the control group were used. These SLRs included articles published up to June 2013. They both included data from MEDLINE and EMBASE , and the one concerning NSAIDs was also performed in Cochrane, DARE (Database of Abstracts of review of Effects), HTA (Health Technology Assessment) database, clinicaltrials.gov, WHO (World Health Organization), ICTRP (International Clinical Trials Registry Platform) and in the websites of the regulatory agencies (e.g. the US Food and Drug Administration (FDA) MedWatch (http://www.fda.gov/Safety/MedWatch/default.htm), the European Medicines Evaluation Agency (http://www.ema.europa.eu), the Australian Adverse Drug Reactions Bulletin (http://www.tga.gov.au/safety/ews-monitoring.htm), and the UK Medicines and Healthcare products Regulatory Agency pharmacovigilance and drug safety updates (http://www.mhra.gov.uk/Safetyinformation/index.htm). These data were updated up to May 2014 using PUBMED, EMBASE, Cochrane, clinicaltrials.gov, and congress abstracts from EULAR and ACR (2012-2014). The language was limited to English, German, French and Spanish. For the data from June 2013 to May 2014, the following keywords were used : ("Spondylitis, Ankylosing" [Mesh] OR ankylosing spondylitis [tw]) AND (clinical trials [tw] OR clinical trial [tw] OR "Clinical Trial" [Publication Type])for PUBMED, and ('ankylosing spondylitis'/exp/mj OR 'ankylosing spondylitis' AND ('disease exacerbation'/exp OR 'disease exacerbation') AND ('clinical trial'/exp OR 'clinical trial') for EMBASE, and (ankylosing spondylitis) for clinical trials.gov and Cochrane.
Data Sources and Searches step 2: Flares-centered studies	Then, all the articles concerning flares in axial SpA in PUBMED and EMBASE were reviewed, with no limit of dates. The language was limited to English, German, French and Spanish. For this step we used the following key words: The language was limited to English, German, French and Spanish. The following key words were used: ("Spondylitis, Ankylosing" [Mesh] OR ankylosing spondylitis [tw]) AND (flares [tw] OR flare [tw] OR exacerbation [tw] OR relapse [tw] OR recurrence [Mesh] OR recurrence [tw] OR clinical reactivation [tw]) for PUBMED , and we used successively the 2 list of key words for EMBASE: ('ankylosing spondylitis'/exp/mj AND ('disease exacerbation'/exp OR 'flare' OR 'flares' OR 'relapse'/exp OR 'recurrent disease'/exp) NOT 'clinical trial'/exp), and ('ankylosing spondylitis'/exp/mj OR 'ankylosing spondylitis'/exp OR

	'ankylosing spondylitis' AND ('disease exacerbation'/exp OR 'disease exacerbation' OR 'relapse'/exp OR 'relapse' OR 'recurrence'/exp OR 'recurrence').
Study selection	For data collection and analysis, one author (AD) assessed independently each title and abstract for suitability for inclusion in the review. If there was any doubt, the full text article was retrieved. Articles were included only if there were RCTs or flares-centered studies with a definition of flare. To be included, RCTs had to concern patients 16 years of age or older with axial SpA whatever the diagnosis/classification criteria. Publications concerning trials already published and included, were excluded from analysis. Studies with the term flares but without definitions of flares, or without full text (e.g. only abstract) or duplicate articles (e.g. already included in the 1st step of SLR) were excluded.
Data extraction	For data extraction, one author (AD) extracted general data and specific data for flares. General data were study identification (first author, journal, year of publication), study characteristics (RCT or other trial intervention, control group, geographical area, sample size, duration of follow-up, criteria for axSpA diagnosis), patient characteristics (average age of patients, percentage of women, percentage of HLA-B27-positive patients, mean disease duration of AxSpA (years since diagnosis), percentage of patients with New York modified criteria, mean BASDAI at baseline). Specific flare data were: the flare design (flare design trial or flare discontinuation trial), the term used for flare (e.g. relapse, recurrence, exacerbation), the origin of the definition (arbitrary or consensual), and the exact definition for flare with: the instrument used, the cutoff, if it was a combination of several instruments or one instrument only, and if it was a relative or absolute change or an absolute value. Then, if possible the number of patients concerned by the definition of flare in the trial was extracted.
Data synthesis and analysis	Analysis was descriptive and included the instrument used to define flare, use of one instrument or of a combination, cutoff used to determine flare, use of a relative or absolute change or use of an absolute value. All the data concerning definitions of flares were classified: word used for “flare”, instrument used to define flare, use of one instrument or of a combination, cutoff used to determine flare, use of a relative or absolute change or an absolute value.

1: Kroon FP, van der Burg LR, Ramiro S, Landewé RB, Buchbinder R, Falzon L, et al. Non-steroidal anti-inflammatory drugs (NSAIDs) for axial spondyloarthritis (ankylosing spondylitis and non-radiographic axial spondyloarthritis). Cochrane Database Syst Rev. 2015;7:CD010952.

2: Bautista-Molano W, Navarro-Compán V, Landewé RBM, Boers M, Kirkham JJ, van der Heijde D. How well are the ASAS/OMERACT Core Outcome Sets for Ankylosing Spondylitis implemented in randomized clinical trials? A systematic literature review. Clin Rheumatol. 2014 Sep;33(9):1313–22.

Online supplementary Table 2: the 38 manuscripts (with the 38 corresponding references) with the definition of flares used and where available, the % of flares in the study

1st author (year) [REF]	Type of study	Drug assessed/ control	Flare design / Flare discontinuation	Sample size (randomized patients)	% males	Age, years (Mean)	Timeframe of flare definition	Definition of flare used	% patients in flare
Mena (1977, South med J) [1]	RCT	NSAID / NSAID	Flare design	26	NA	NA	6 weeks	clear increase in spinal or SI pain and ≥1 of the criteria: muscle spasm in back/decreased range of motion of some part of the spine/increased ESR	NA
Mena (1977) [2]	RCT	NSAID/ phenylbutazone	Flare design	26	NA	NA	6 weeks	clear increase in spinal or SI pain and ≥1 of the criteria: muscle spasm in back/decreased range of motion of some part of the spine/increased ESR	NA
Ansell (1978) [3]	RCT	NSAID / NSAID	Flare design	25	8%	NA	2 weeks	deterioration in at least 2 criteria among: morning stiffness, immobility stiffness, morning pain,	NA

								discomfort, Schober, wall tragus distance	
Bryon (1982) [4]	RCT	NSAID / tolmetin sodium	Flare design	34	22%	43.6	4 weeks	deterioration of at least 2 of these criteria: pain on spinal palpation/morning stiffness/nocturnal pain/subjective pain/immobility stiffness	86%
Fransen (1986) [5]	RCT	NSAID/ phenylbutazone	Flare design	246	0%	37	14 days	a worsening of patient's condition in which pain and stiffness was an essential component requiring treatment	NA
Khan (1987) [6]	RCT	NSAID/ NSAID	Flare design	262	16%	NA [19-67]	2-15 days	2 criteria among: 1 point increase in cervical/thoracic/lumbar/SI pain on a 0-4 point scale assessed by patient and ≥2 of criteria: increased duration of morning stiffness (≥30 min)/decreased shober (≥1 cm)/decreased chest expansion (≥1 cm)/increased	NA

								distance fingertips to floor (≥5cm)/≥1 periph joint affected by swelling and tenderness/ESR ≥28mm	
Schwarzer (1990) [7]	RCT	NSAID /NSAID	Flare design	24	12%	41	3 days	increase in back pain and stiffness	NA
Dougados (1994) [8]	RCT	NSAID / placebo	Flare design	285	26%	40	2 days	pain > 40mm on VAS 100mm and increase in pain of at least 30% between the screening and the entry visit	90%
Dougados (1999) [9]	RCT	NSAID / placebo	Flare design	473	22%	32.5	2-15 days	pain ≥40mm on VAS 100mm and an increase at least 30% between the screening visit	91%
Ruof (1999) [10]	Non randomized controlled trial	NSAID / vitamin / placebo	Flare discontinuation	120	25%	44	NA	predefined magnitude and duration of deterioration in back pain	NA
Dougados (2001) [11]	RCT	NSAID / placebo	Flare design	246	31%	38.6	14 days	pain scored ≥40 mm on VAS 100mm and an increase of at least 30% between the screening and the baseline visit	31%

Breban (2002) [12]	Open study	antiTNF	Flare discontinuation	50	24%	35	NA	≥50% loss of global assessment of pain	NA
Baraliakos (2005) (outcome…) [13]	Open extension study	Anti TNF	Flare discontinuation	26	23%	37.1	NA	BASDAI ≥4 and physician global assessment ≥4	NA
Baraliakos (2005)(clinical response…) [14]	Open extension trial	Anti TNF	Flare discontinuation	42	NA	NA	NA	BASDAI ≥4 and physician global assessment ≥4	NA
Brandt (2005) [15]	Open observational study	Anti TNF	Flare discontinuation	26	23%	37.1	NA	BASDAI ≥4 and physician global assessment ≥4	NA
Van der Heijde (2005) [16]	RCT	NSAID / NSAID	Flare design	387	22%	43.6	NA	worsening of AS defined as ≥40mm on patient's assessment of spine pain and an increase of ≥30% (min 12mm)compared with screening period	NA
Wanders (2005) [17]	RCT	NSAID / NSAID	Flare design	215	31%	38.5	NA	pain≥ 40mm and increase ≥30%	NA
Barkhuizen (2006) [18]	RCT	NSAID/ placebo	Flare design	611	26%	44.5	2 weeks	pain on VAS≥50 mm worsening by 30% compared with the	NA

								preinclusion visit	
Boonen (2006) [19]	Observational	antiTNF	Cost utility	130	29%	45.9	NA	BASDAI≥4	NA
Baraliakos (2007) [20]	Extension study	Anti TNF	Flare discontinuation	NA	NA	NA	NA	BASDAI ≥4 and physician global assessment ≥4	NA
Huang (2007) [21]	Open study	Anti TNF	Flare discontinuation	63	37%	32.8	NA	BASDAI >60% of the corresponding score at baseline	NA
Breban (2008) [22]	RCT	Anti TNF/ anti TNF	Flare discontinuation	247	24%	41	NA	a negative answer to the 1st question ("since the last connection, did you think that your disease has remained under control?") and a positive answer to the 2nd question ("since…, do you think that your disease has been worsening?") and either an increase in BASDAI score of ≥1/10 or an increase in patient's assessment of ≥2/10 compared with the lowest score reached by the patient since	NA

								the 1st infliximab infusion	
Sieper (2008) [23]	RCT	NSAID/ NSAID	Flare design	458	31%	44.8	2-14 days	pain≥ 40mm and increase ≥30%	NA
Krzysiek (2009) [24]	RCT	antiTNF	Flare discontinuation	169	23%	40	NA	negative response at "since the last connection, do you thinck that your disease has remained under control?"AND a positive question at ""do you think that your disease has been worsening?" AND either an increase in BASDAI score ≥1 or an increase in global pain score ≥2 as compared with the lowest score reached by that patient since the 1st infusion.	NA
Baraliakos (2010) [25]	review	Anti TNF	NA	NA	NA	NA	NA	BASDAI ≥4 and physician global assessment ≥4	NA
Maksymowych (2010) [26]	Not interventional study	NA	NA	291	25%	45.7		2 questions: “Are you currently experiencing a flare of your AS?” and “Is your AS	49%

								sufficiently active to require an assessment and examination from your rheumatologist?”	
Heldmann (2011) [27]	Open study	Anti TNF	Flare discontinuation	103	17%	44	NA	BASDAI> 4 and physician's global assessment > 4 at screening and baseline	NA
Gratacós Jordi (2012) [28]	RCT	AntiTNF	Flare discontinuation	NA	NA	NA	NA	BASDAI > 4, global clinical impression by physician >4 and at least one of three following criteria: patient impression >= 4, axial nocturnal pain (VAS) >= 4, and increased of acute phase reactants (reactive C protein (PCR) and/or erythrocyte sedimentation rate (ESR))	NA
Song (2012) [29]	Open study	Anti TNF	Flare discontinuation	17	29%	NA	NA	BASDAI increase of 2 points	NA
Cantini (2013) [30]	RCT	Anti TNF/ anti TNF	Flare discontinuati	78	28%	38	NA	BASDAI > 4 or any of the other above-	NA

			on					mentioned peripheral articular and extra-articular manifestations independently of elevation of acute-phase reactants	
Deng (2013) [31]	RCT	Thalidomide /SLZ / NSAID	Flare design	111	NA	NA	NA	increase of BASDAI ≥2 or BASDAI degradation≥80%	NA
Haibel (2013) [32]	RCT	AntiTNF	Flare discontinuation	24	54%	37.5	NA	loss of an established ASAS40 response as compared to baseline at any timepoint	NA
Jaclyn K Anderson (2013) [33]	RCT	Anti TNF/ placebo	Flare discontinuation	NA	NA	NA	NA	2 consecutive study visits with ASDAS ≥ 2.1	NA
Song (2013) [34]	Open study	Rituximab	Flare discontinuation	9	NA	NA	NA	1,5 point worsening of the BASDAI compared to the lowest BASDAI	NA
Kadar (2014) [35]	Retrospective	NA	NA	17	NA	46	Between 2 visits	when the disease activity assessed by BASDAI became high from low or moderate activity at the previous visit	50%
Sieper (2014)	RCT	Anti TNF/	Flare design	158	28%	31.4	NA	increase in total back	NA

(INFAST part1) [36]		placebo						pain≥30%	
Sieper (2014) (INFAST part 2) [37]	RCT	NSAID/ placebo	Flare design	82	22%	29	2 consecutive visits within 1-3 weeks of each other	BASDAI (0-10) ≥3 cm	7%
Dougados (2014) [38]	RCT	AntiTNF/placebo	Flare design	90	38%	39	NA	Symptom flare	70%

References

1. Mena HR, Good AE. Management of ankylosing spondylitis with flurbiprofen or indomethacin. South Med J. 1977 Aug;70(8):945-7.
2. Mena HR, Willkens RF. Treatment of ankylosing spondylitis with flurbiprofen or phenylbutazone. European Journal of Clinical Pharmacology 1977;11(4):263-266.
3. Ansell BM, Major G, Liyanage SP, Gumpel JM, Seifert MH, Mathews JA, Engler C. A comparative study of Butacote and Naprosyn in ankylosing spondylitis. Annals of the Rheumatic Diseases 1978;37(5):436-9.
4. Bryon MA, Steele EC. A double-blind crossover study comparing tolmetin sodium and naproxen in the treatment of ankylosing spondylitis. Curr Med Res Opin. 1982; 7(10): 670-676.
5. Franssen MJ, Gribnau FW, van de Putte LB. A comparison of diflunisal and phenylbutazone in the treatment of ankylosing spondylitis. Clinical Rheumatology 1986;5(2):210-220.
6. Khan MA. A double blind comparison of diclofenac and indomethacin in the treatment of ankylosing spondylitis. Journal of Rheumatology 1987;14(1):118-123.
7. Schwarzer AC, Cohen M, Arnold MH, Kelly D, McNaught P, Brooks PM. Tenoxicam compared with diclofenac in patients with ankylosing spondylitis. Current Medical Research & Opinion 1990;11(10):648-653.

8. Dougados M, Nguyen M, Caporal R, Legeais J, Bouxin-Sauzet A, Pellegri-Guegnault B, Gomeni C. Ximoprofen in ankylosing spondylitis. A double blind placebo controlled dose ranging study. Scandinavian Journal of Rheumatology 1994;23(5):243-248.

9. Dougados M, Gueguen A, Nakache JP, Velicitat P, Veys EM, Zeidler H, Calin A. Ankylosing spondylitis: what is the optimum duration of a clinical study? A one year versus a 6 weeks non-steroidal anti-inflammatory drug trial. Rheumatology 1999;38(3):235-44.

10. Ruof J, Sangha O, Stucki G. Comparative responsiveness of 3 functional indices in ankylosing spondylitis. J Rheumatol. 1999 Sep;26(9):1959–63.

11. Dougados M, Behier JM, Jolchine I, Calin A, van der Heijde D, Olivieri I, Zeidler H, Herman H. Efficacy of celecoxib, a cyclooxygenase 2-specific inhibitor, in the treatment of ankylosing spondylitis: a six-week controlled study with comparison against placebo and against a conventional nonsteroidal antiinflammatory drug. Arthritis & Rheumatism 2001;44(1):180-185.

12. Breban M, Ravaud P, Claudepierre P, Baron G, Henry YD, Hudry C, Euller-Ziegler L, Pham T, Solau-Gervais E, Chary-Valckenaere I, Marcelli C, Perdriger A, Le Loët X, Wendling D, Fautrel B, Fournié B, Combe B, Gaudin P, Jousse S, Mariette X, Baleydier A, Trape G, Dougados M; French Ankylosing Spondylitis Infliximab Network. Maintenance of infliximab treatment in ankylosing spondylitis: results of a one-year randomized controlled trial comparing systematic versus on-demand treatment. Arthritis Rheum. 2008 Jan;58(1):88-97.

13. Baraliakos X, Listing J, Brandt J, Zink A, Alten R, Burmester G, et al. Clinical response to discontinuation of anti-TNF therapy in patients with ankylosing spondylitis after 3 years of continuous treatment with infliximab. Arthritis Res Ther. 2005;7(3):R439–444.

14. Baraliakos X, Brandt J, Listing J, Haibel H, Sörensen H, Rudwaleit M, et al. Outcome of patients with active ankylosing spondylitis after two years of therapy with etanercept: clinical and magnetic resonance imaging data. Arthritis Rheum. 2005 Dec 15;53(6):856–63.

15. Brandt J, Listing J, Haibel H, Sörensen H, Schwebig A, Rudwaleit M, et al. Long-term efficacy and safety of etanercept after readministration in patients with active ankylosing spondylitis. Rheumatology (Oxford). 2005 Mar;44(3):342–8.

16. van der Heijde D, Baraf HSB, Ramos-Remus C, Calin A, Weaver AL, Schiff M, Jamers M, Markind JE, Reicin AS, Melian A, Dougados M. Evaluation of the efficacy of etoricoxib in ankylosing spondylitis: results of a fifty-two-week, randomized, controlled study. Arthritis & Rheumatism 2005;52(4):1205-1215.

17. Wanders A, van der Heijde D, Landewe R, Behier JM, Calin A, Olivieri I, Zeidler H, Dougados M. Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis & Rheumatism 2005;52(6):1756-1765.

18. Barkhuizen A, Steinfeld S, Robbins J, West C, Coombs J, Zwillich S. Celecoxib is efficacious and well tolerated in treating signs and symptoms of ankylosing spondylitis. Journal of Rheumatology 2006;33(9):1805-1812.

19. Boonen A, van der Heijde D, Severens JL, Boendermaker A, Landewé R, Braun J, et al. Markov model into the cost-utility over five years of etanercept and infliximab compared with usual care in patients with active ankylosing spondylitis. Ann Rheum Dis. 2006 Feb;65(2):201–8.

20. Baraliakos X, Listing J, Rudwaleit M, Brandt J, Alten R, Burmester G, et al. Safety and efficacy of readministration of infliximab after longterm continuous therapy and withdrawal in patients with ankylosing spondylitis. J Rheumatol. 2007 Mar;34(3):510–5.

21. Huang F, Zhu J, Zhang L, Zhang J, Zhang F, Yu D. Response to one infusion predicts subsequent improvement as well as the rate of relapse of ankylosing spondylitis infused with three pulses of infliximab. Clin Rheumatol. 2007 Jun;26(6):920–6.

22. Breban M, Ravaud P, Claudepierre P, Baron G, Henry YD, Hudry C, Euller-Ziegler L, Pham T, Solau-Gervais E, Chary-Valckenaere I, Marcelli C, Perdriger A, Le Loët X, Wendling D, Fautrel B, Fournié B, Combe B, Gaudin P, Jousse S, Mariette X, Baleydier A, Trape G, Dougados M; French Ankylosing Spondylitis Infliximab Network. Maintenance of infliximab treatment in ankylosing spondylitis: results of a one-year randomized controlled trial comparing systematic versus on-demand treatment. Arthritis Rheum. 2008 Jan;58(1):88-97.

23. Sieper J, Klopsch T, Richter M, Kapelle A, Rudwaleit M, Schwank S, Regourd E, May M. Comparison of two different dosages of celecoxib with diclofenac for the treatment of active ankylosing spondylitis: results of a 12-week randomised, double-blind, controlled study. Ann Rheum Dis. 2008 Mar;67(3):323-9

24. Krzysiek R, Breban M, Ravaud P, Prejean MV, Wijdenes J, Roy C, et al. Circulating concentration of infliximab and response to treatment in ankylosing spondylitis: results from a randomized control study. Arthritis Rheum. 2009 May 15;61(5):569–76.

25. Baraliakos X, Braun J. Anti-TNF-alpha therapy with infliximab in spondyloarthritides. Expert Rev Clin Immunol. 2010 Jan;6(1):9–19.

26. Heldmann F, Brandt J, van der Horst-Bruinsma IE, Landewe R, Sieper J, Burmester GR, et al. The European ankylosing spondylitis infliximab cohort (EASIC): a European multicentre study of long term outcomes in patients with ankylosing spondylitis treated with infliximab. Clin Exp Rheumatol. 2011 Aug;29(4):672–80

27. Maksymowych WP, Richardson R, Mallon C, van der Heijde D, Boonen A. Evaluation and validation of the patient acceptable symptom state (PASS) in patients with ankylosing spondylitis. Arthritis Rheum. 2007 Feb 15;57(1):133-9.

28. (189) NCT 01604629. Gratacós Jordi Evaluation of a Standardized Protocol for Dose Reduction in Patients With Spondylarthropathies and Clinical Remission With Anti-TNF Therapy.

29. Song I-H, Althoff CE, Haibel H, Hermann K-GA, Poddubnyy D, Listing J, et al. Frequency and duration of drug-free remission after 1 year of treatment with etanercept versus sulfasalazine in early axial spondyloarthritis: 2 year data of the ESTHER trial. Ann Rheum Dis. 2012 Jul;71(7):1212–5.

30. Cantini F , Niccoli L , Cassara E , Kaloudi O and Nannini C . Duration of remission after halving of the etanercept dose in patients with ankylosing spondylitis : A randomized, prospective, long-term, follow-up study. Biologics: Targets and Therapy, 2013, 7(1), 1.

31. Deng X, Zhang J, Zhang J, Huang F. Thalidomide reduces recurrence of ankylosing spondylitis in patients following discontinuation of etanercept. Rheumatol Int. 2013 Jun;33(6):1409-13.

32. Haibel H, Heldmann F, Braun J, Listing J, Kupper H, Sieper J. Long-term efficacy of adalimumab after drug withdrawal and retreatment in patients with active non-radiographically evident axial spondyloarthritis who experience a flare. Arthritis Rheum. 2013 Aug;65(8):2211-3.

33. (185) NCT01808118. Jaclyn K Anderson Study Comparing the Efficacy and Safety of Continuing Versus Withdrawing Adalimumab Therapy in Maintaining Remission in Subjects With Non-Radiographic Axial Spondyloarthritis.

34. Song I-H, Heldmann F, Rudwaleit M, Listing J, Appel H, Haug-Rost I, et al. One-year follow-up of ankylosing spondylitis patients responding to rituximab treatment and re-treated in case of a flare. Ann Rheum Dis. 2013 Feb;72(2):305–6.

35. Kádár G, Balázs E, Soós B, Laduver A, Keszthelyi P, Szekanecz Z, et al. Disease activity after the discontinuation of biological therapy in inflammatory rheumatic diseases. Clin Rheumatol. 2014 Mar;33(3):329–33.

36. Sieper J, Lenaerts J, Wollenhaupt J, Rudwaleit M, Mazurov VI et al. Efficacy and safety of infliximab plus naproxen versus naproxen alone in patients with early, active axial spondyloarthritis: results from the double-blind, placebo-controlled INFAST study, Part 1. Ann Rheum Dis. 2014 Jan;73(1):101-7

37. Sieper J, Lenaerts J, Wollenhaupt J, Rudwaleit M, Mazurov V. Maintenance of biologic-free remission with naproxen or no treatment in patients with early, active axial spondyloarthritis: results from a 6-month, randomised, open-label follow-up study, INFAST Part 2. Ann Rheum Dis. 2014 Jan;73(1):108-13

38. Dougados M, Wood E, Combe B, Schaeverbeke T, Miceli-Richard C, Berenbaum F, Koppiker N, Dubanchet A, Logeart I. Evaluation of the nonsteroidal anti-inflammatory drug-sparing effect of etanercept in axial spondyloarthritis: results of the multicenter, randomized, double-blind, placebo-controlled SPARSE study. Arthritis Res Ther. 2014 Nov 27;16(6):481.

Online supplementary Table 3. The methods used in the vignette exercise and examples of vignettes

Development of the case-vignettes	The case-vignettes were elaborated by 3 authors (AD, LG and MD) during a meeting (July 16th 2014). A vignette is a "brief written case history of a fictitious patient based on a realistic clinical situation that is accompanied by 1 or more questions that explore what a physician would do if presented with the actual patient". (Veloski 2005). Vignettes should be realistic, brief and precise. It was decided to have the vignettes only in English.
Scenario	In this study, only one scenario was used for all the vignettes. "A 32 year-old man with a well-established diagnosis of axial SpA consults you at 2 successive timepoints. In comparison to the previous visit and according to the following data, and all other things being equal (physical examination, CRP and NSAID intake), do you consider this patient is flaring at the second visit? Yes or No. Please give an answer (yes or no) even if you are unsure." It was decided that a single phenotype of axial SpA would be proposed in the scenario, excluding potential signs of fibromyalgia or peripheral SpA, in order to avoid an excessive number of vignettes, and considering that defining flares in fibromyalgia is not our objective. The disease duration was not given since we aimed for a flare definition which would not depend on disease duration. The timeframe between visits was not determined to allow better external validity of the definition.
Variations in the vignettes	For each case-vignette, different values of instruments used to define flares were proposed by the authors. It was decided that vignettes would describe variations of one instrument only, and not a combination of variations of several instruments, i.e., "all other things being equal". The vignettes did not give any information other than the variable of interest. For each question, participants chose the "yes" or "no" answer. No other answer was possible.
Selection of relevant domains for vignettes	Instruments used in case-vignettes to assess flares in axial SpA were chosen as: ✓ BASDAI on a numeric rating scale (0-10) ✓ Pain due to axial SpA on a numeric rating scale (0-10) ✓ ASDAS score ✓ Initially, CRP and NSAID intake BASDAI and pain were selected because there were the 2 most frequent instruments in the literature used to define flares in axial SpA. ASDAS score was selected, because this is a recent instrument validated in axial SpA by ASAS. CRP was selected because a number of studies used this instrument to assess flares in axial SpA.

	NSAID intake was selected, because NSAIDs represent a treatment of short term efficacy useful in daily practice when a disease exacerbation exists. Thus, an increase of NSAIDs intake could reflect a flare of axial SpA.
Variations of outcomes values	It was decided that BASDAI and pain could vary only by a minimum of 1/10 on NRS (Numeric Rating Scale). Minimal initial BASDAI and initial pain were 0, and maximal initial BASDAI and pain were 6 (higher baseline values were excluded, because referring to a non-controlled disease). Variations of ASDAS were decided taking into account minimal clinically important differences for this score. Variations of CRP were arbitrarily decided with steps of 5, 10, 20, 30, with initial values of: <5, 8, 13, 18, and 25 mg/l. NSAIDs intake was assessed using one NSAID, Naproxen, with no intake or intake of half dose (500mg per day) or full dose (1000mg per day). Initial frequencies of intake were: no intake, intake <1 day/ week, 1 to 3 days/week, 3 to 5 days/week, ≥5 days, and only for the 500mg dosage: everyday. Final frequencies of intake could be: intake <1 day/week, 1 to 3 days/week, 3 to 5 days/week, ≥5 days/week, every day (7 days/week).
Sending out the vignettes to the ASAS members	All the 159 ASAS experts were asked to assess a sample of 46 vignettes between July and December 2014; each sample was intentionally constructed to include vignettes for each outcome and a distribution of changes in status. The choice of vignettes' attribution was different for each ASAS member, and comprised 5 vignettes concerning each outcome for the first round, and 21 additional vignettes for the second round. The attributions were made using excel and included a good spread between different variations for each participant (to make sure one person did not receive only vignettes with changes of one point for each outcome for example).
Covering email and instructions used for the vignette exercise	Dear ASAS member, This email is to ask for your participation in an ASAS project. This ASAS project aims to find a consensual definition of flare in axial SpA to be used in the context of clinical trials and longitudinal studies. This project is led by Maxime Dougados and is endorsed by the ASAS Executive. We remind you it was presented during the ASAS sessions last January and then in June during the EULAR congress. As an ASAS member, we count on your participation and we kindly remind you that participation in ASAS initiatives is part of being an ASAS member. For this initiative, you are invited to answer to case-vignettes concerning flares in axial SpA. You have been attributed 25 case-vignettes for which your answer is required as 'flare yes/no'. The exercise is very quick; it will take you no more than 5-10 minutes to answer. Please find attached an Excel file including the scenario and different vignettes. Please save the file to your computer, complete the column about flares yes/no and please send it back to us or use 'reply' to this email. Please can we ask you to send your answer within 2 weeks.

	Thank you for your participation, Best wishes There were no instructions regarding what should be defined as a “flare”. The document opened onto this text : for this initiative, you are invited to answer to case-vignettes concerning flares in axial SpA. You have been attributed 25 case-vignettes for which your answer is required as ‘flare yes/no’. Please select in each of these cases if the patient is flaring at the second visit, yes or no.
Vignette example for Pain	A 32 year-old man with a well-established diagnosis of axial SpA consults you at 2 successive timepoints. In comparison to the previous visit and according to the following data, and all other things being equal (physical examination, CRP and NSAID intake), do you consider this patient is flaring at the second visit? Yes or No. Please give an answer (yes or no) even if you are unsure. - Initial (first visit) Pain score (0-10 NRS): 3 - Final (second visit) Pain score (0-10 NRS): 7 - Flare: Yes/No Reminder. Pain due to axial SpA is assessed as follows: "Circle the number between 0 and 10 that best describes the pain you felt due to spondyloarthritis during the last 48 hours". Interpretation: Pain levels below 4 are usually considered acceptable.
Vignette example for BASDAI	A 32 year-old man with a well-established diagnosis of axial SpA consults you at 2 successive timepoints. In comparison to the previous visit and according to the following data, and all other things being equal (physical examination, CRP and NSAID intake), do you consider this patient is flaring at the second visit? Yes or No. Please give an answer (yes or no) even if you are unsure. - Initial (first visit) BASDAI (0-10): 2 - Final (second visit) BASDAI (0-10): 4 - Flare: Yes/No Reminder. The BASDAI takes into account axial, peripheral and entheseal pain, fatigue and morning stiffness. Interpretation: an active disease is defined by a BASDAI ≥4/10, and an improvement of 2 points on BASDAI is the minimal clinically important improvement. There is no definition of the minimal clinically important worsening/deterioration.

Vignette example for ASDAS-CRP	A 32 year-old man with a well-established diagnosis of axial SpA consults you at 2 successive timepoints. In comparison to the previous visit and according to the following data, and all other things being equal (physical examination, CRP and NSAID intake), do you consider this patient is flaring? Yes or No. Please give an answer (yes or no) even if you are unsure - Initial ASDAS score: 0.8 - Final ASDAS score : 2.3 - Flare: Yes/No Reminder. The ASDAS score takes into account back pain, duration of morning stiffness, patient global assessment, peripheral pain/ swelling and CRP. Interpretation: an ASDAS score<1.3 defines an inactive disease, an ASDAS score between 1.3 and 2.1 corresponds to a moderate disease activity, an ASDAS score between 2.1 and 3.5 reflects a high disease activity, and an ASDAS sore above 3.5 corresponds to a very high disease activity. An improvement of 1.1 on ASDAS score defines a clinically important improvement. There is no definition of the minimal clinically important worsening/deterioration.

Supplementary online figure 1. The ROC curves for each outcome to determine flares.

Figure A. ROC curve to define flares based on changes in pain

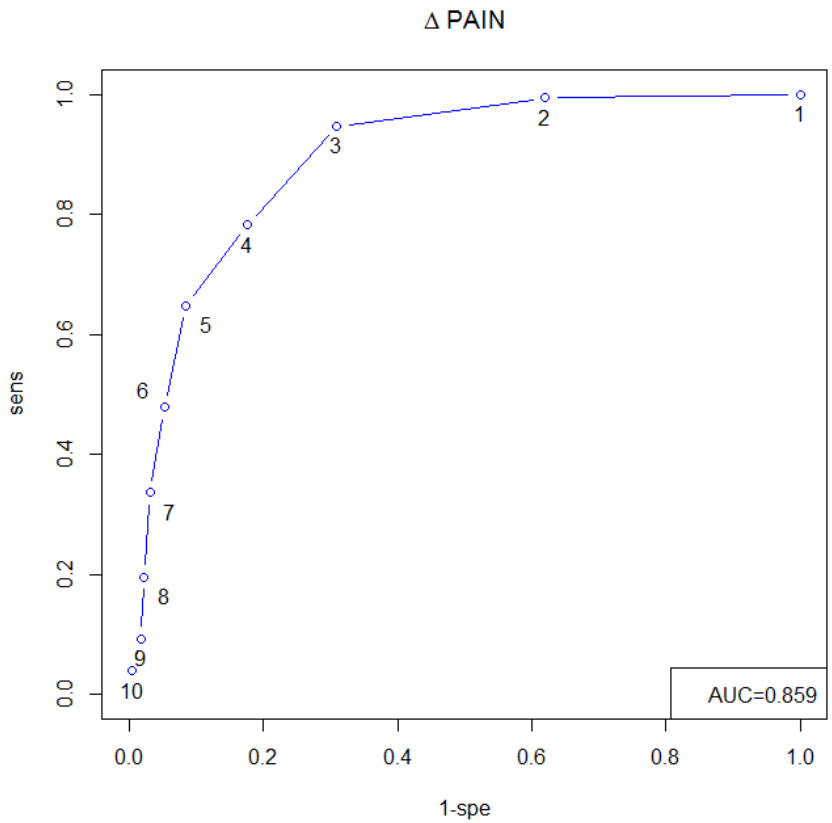


Figure B. ROC curve to define flares based on changes in BASDAI

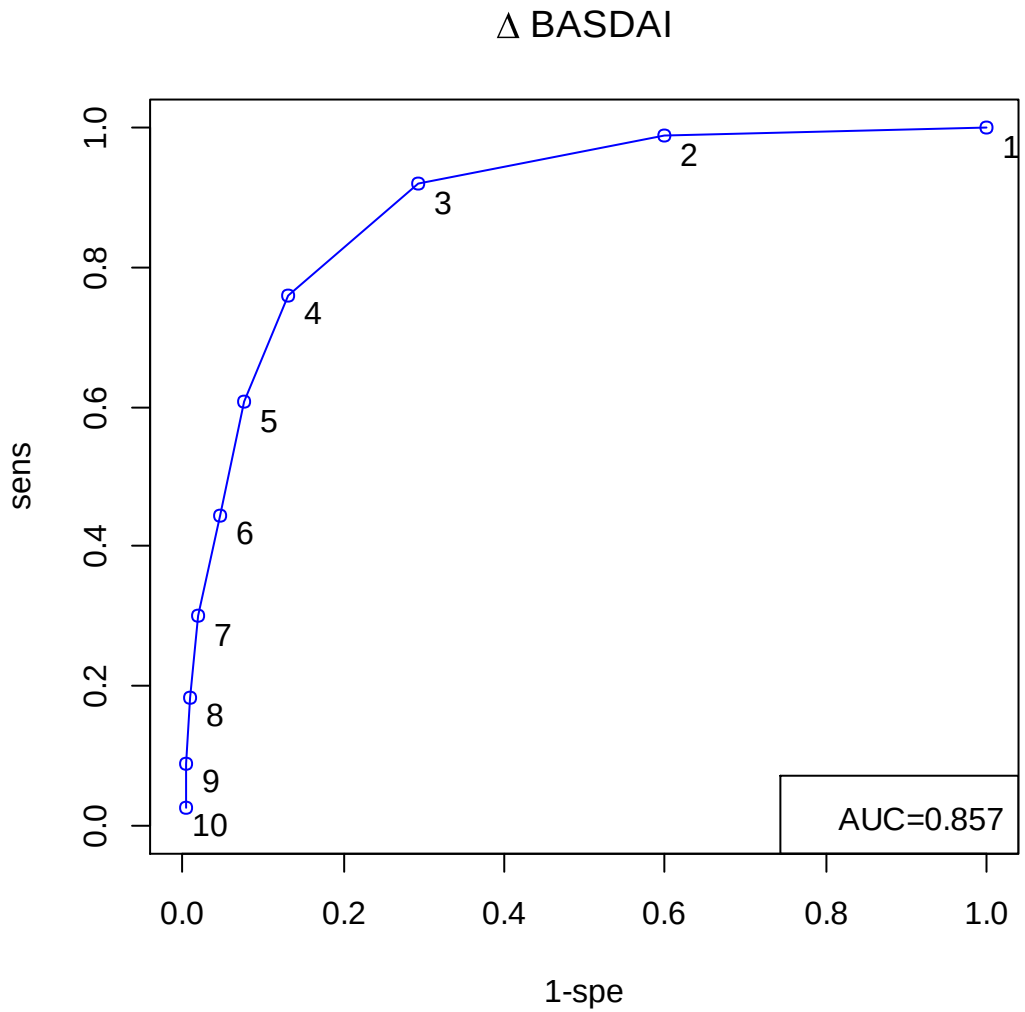
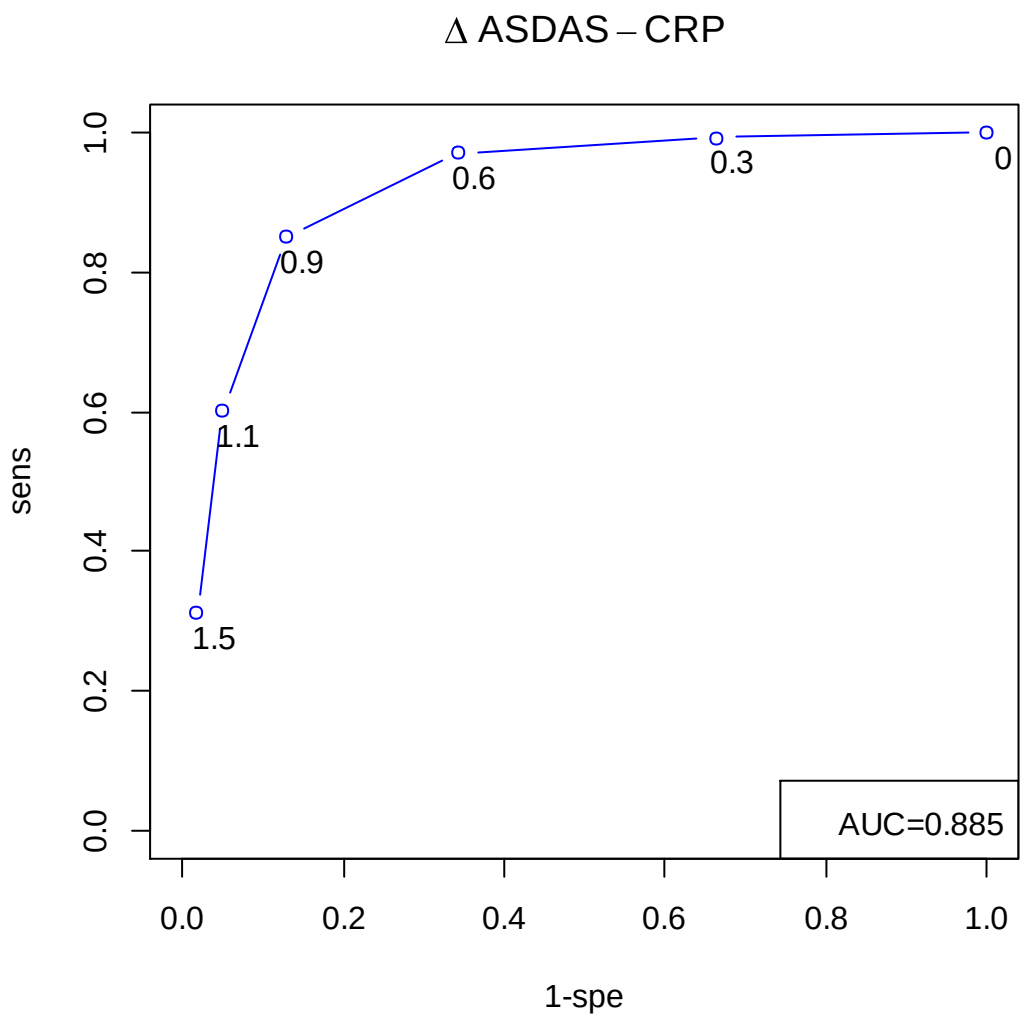


Figure C. ROC curve to define flares based on changes in ASDAS CRP



Annals of the Rheumatic Diseases
The EULAR Journal
BMJ
ard.bmj.com

A step closer to defining 'flare' in people with axial spondyloarthritis

Flares may be used as outcomes in trials and studies, but to date there is no agreed consensus on how to define them in people with axial spondyloarthritis.

INTRODUCTION

Spondyloarthritis is an umbrella term for several conditions that share many of the same features and symptoms, including ankylosing spondylitis, psoriatic arthritis and reactive arthritis. Patients can also be classified as having axial or non-axial (peripheral) disease, according to which joints in their body are affected. Axial disease affects the sacroiliac joint (in the back part of the pelvis) causing back pain and stiffness.

People with axial spondyloarthritis may suffer from flares of their disease, when the symptoms get much worse. These flares tend to alternate with periods of low disease activity or remission, when a person may feel well.

WHAT DID THE AUTHORS HOPE TO FIND?

The authors wanted to work out how best to define flares (disease worsening) with a number of commonly used tools for measuring disease activity in people with axial spondyloarthritis.

WHO WAS STUDIED?

No real people were used in this study.

HOW WAS THE STUDY CONDUCTED?

This was a systematic literature review, which means that the authors identified all the published evidence on a particular topic and drew it together into one summary. They then performed a second exercise where they used made-up patient cases to see how well experts in axial spondyloarthritis agreed on the definition of a flare. In all, 121 expert doctors looked at 46 patient cases and reported whether each one was in flare or not.

WHAT WERE THE MAIN FINDINGS OF THE STUDY?

The main finding was that there is no real agreement about how to define a flare in axial spondyloarthritis, even between experts in the disease. Typically, most doctors would say that a change of around 3 points on a scale from 0 to 10 is a flare.

ARE THESE FINDINGS NEW?

Yes, this is the first time that anyone has tried to define a flare in axial spondyloarthritis.

WHAT ARE THE LIMITATIONS OF THE STUDY?

One main limitation is that the study did not look at real people with axial spondyloarthritis. Additionally, it did not ask real patients with the disease what they consider to be a flare.

WHAT DO THE AUTHORS PLAN ON DOING WITH THIS INFORMATION?

More work is being done in this area to help to define flares better. A second study is underway looking at real-life people with axial spondyloarthritis. This will ask them to assess their own disease and flag when they think they are having a flare.

WHAT DOES THIS MEAN FOR ME?

If you have axial spondyloarthritis, these studies might mean that in the future your doctor will be better able to talk to you about flares and how to define them. This may also help to standardise the way you and other people with the disease are treated. If you are concerned about your disease or your symptoms, you should talk to your doctor.

professionals for all aspects of their health care and only rely on the Summary if directed to do so by their medical professional. Please view our full Website Terms and Conditions. http://www.bmj.com/company/legal-information/

Date prepared: June 2016

Summary based on research article published on: 4 February 2016

From: Gossec, L. *et al*. Preliminary definitions of 'flare' in axial spondyloarthritis, based on pain, BASDAI and ASDAS-CRP: an ASAS initiative. *Ann Rheum Dis* 2016;75:991–6. doi:10.1136/annrheumdis-2015-208593