



Editor's choice  
Scan to access more  
free content

## EXTENDED REPORT

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

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

**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-Salpêtrière, Service de Rhumatologie, 47-83, boulevard de l'Hôpital, Paris 75013, France; [laure.gossec@aphp.fr](mailto: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.

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.<sup>3,4</sup>

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.<sup>5–7</sup> 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.<sup>8,9</sup> There are previously published studies on the perception of flare by the patient in SpA.<sup>10–12</sup> 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.<sup>13</sup>

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.

## INTRODUCTION

The natural course of axial spondyloarthritis (axSpA) includes periods of flares and remission.<sup>1</sup> 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)<sup>2</sup>;



CrossMark



Linked

► <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.<sup>14</sup>

### 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<sup>15</sup> 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<sup>16</sup> (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.<sup>16–18</sup>

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  $\geq X$  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  $\geq X$ , 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  $\geq 4/10$  with absolute physician assessment  $\geq 4/10$  used in six studies, and increase in pain  $\geq 30\%$  with absolute pain  $\geq 4/10$  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  $\geq 2$  points and increase in pain  $\geq 3$  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  $\geq 2$  points, more than 70% of the doctors will consider there is a flare if the referral level of pain is  $\leq 4$ . (b) Considering a pain change  $\geq 2$  points, more than 60% of the doctors will consider there is a flare if the final value is  $\geq 4$ . (c) Considering a pain change  $\geq 3$  points, more than 80% of the doctors will consider there is a flare if the referral level of pain is  $\leq 4$ . (d) Considering a pain change  $\geq 3$ , more than 80% of the doctors will consider there is a flare if the final value is  $\geq 5$ .

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  $\geq 2$  and an observed value at the time of the flare of  $\geq 4$ ; (b) an increase in pain of  $\geq 3$ . 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  $\geq 4$ , a 'flare' is defined as an increase in pain  $\geq 2$  points, otherwise, flare is defined as an increase in pain  $\geq 3$  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  $\geq 2$  points and increase in BASDAI  $\geq 3$  points. For these two cut-offs, the performances were again calculated for different referral and observed values. (a) Considering a BASDAI change  $\geq 2$ , more than 80% of the doctors will consider there is a flare if the referral BASDAI is  $\leq 4$ . (b) Considering a BASDAI change  $\geq 2$ , more than 60% (or 70%) of the doctors will consider there is a flare if the observed value is  $\geq 4$  (or 5). (c) Considering a BASDAI change  $\geq 3$ , more than 80% of the doctors will consider there is a flare if the referral BASDAI is  $\leq 4$ . (d) Considering a BASDAI change  $\geq 3$ , more than 70% of the doctors will consider there is a flare if the observed value is  $\geq 4$  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  $\geq 4$ ,

**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 $\geq 4$ (N=9) Abs. value $\geq 3$ (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 $\geq 2.1$ (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 $\geq 40$ mm and increase of 30% (N=7) Abs. value $\geq 40$ 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 $\geq 30$ min (N=1)
	Patient global assessment (0–10)	2 (5.2)	Combination N=2	Abs. value $\geq 4/10$ (N=1) Rel. change $\geq 2/10$ (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 $\geq 4$ (N=7)
	Physical assessment	5 (13.2)	Combination N=5	Restriction (N=4) Abs. change: decreased Schöber index ( $\geq 1$ cm), decreased chest expansion ( $\geq 1$ cm), increased fingertips to floor distance ( $\geq 5$ cm)
	Acute phase reactants	4 (10.4)	Combination N=4	Presence (N=3) ESR $\geq 28$ 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<sup>16</sup>; axSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index<sup>14</sup>; ESR, erythrocyte sedimentation rate; NSAIDs, non-steroidal anti-inflammatory drugs; Rel: relative.

‘flare’ is defined as an increase in BASDAI  $\geq 2$  points; otherwise, ‘flare’ is defined as an increase in BASDAI  $\geq 3$  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  $\geq 2$  points than when the change in BASDAI was  $\geq 3$  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	$\Delta$ pain $\geq 2$ AND final value $\geq 4$	0.99	0.30	0.76	0.97
		$\Delta$ pain $\geq 3$	0.95	0.69	0.83	0.88
		If observed value is $\geq 4$ : $\Delta$ pain $\geq 2$ points, otherwise: $\Delta$ pain $\geq 3$ points	0.97	0.56	0.79	0.92
BASDAI (0–10)	0.86	$\Delta$ BASDAI $\geq 2$ points	0.99	0.40	0.78	0.94
		$\Delta$ BASDAI $\geq 2$ points AND final value $\geq 4$	0.99	0.32	0.81	0.92
		$\Delta$ BASDAI $\geq 3$ points	0.92	0.70	0.87	0.80
		$\Delta$ BASDAI $\geq 3$ points AND final value $\geq 4$	0.94	0.63	0.88	0.79
		If observed value is $\geq 4$ , $\Delta$ BASDAI $\geq 2$ points, otherwise: $\Delta$ BASDAI $\geq 3$ points	0.94	0.54	0.81	0.80
ASDAS-CRP	0.89	$\Delta$ ASDAS $\geq 0.6$	0.97	0.65	0.75	0.96
		$\Delta$ ASDAS $\geq 0.9$	0.85	0.87	0.87	0.85
		$\Delta$ ASDAS $\geq 1.1$	0.60	0.94	0.93	0.69
		$\Delta$ ASDAS $\geq 0.6$ AND observed ASDAS $\geq 1.3$	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;  $\Delta$ : 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  $\geq 0.6$ , 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  $\geq 1.3$  (ie, not being in inactive disease<sup>18</sup>) 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.<sup>19 20</sup>

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.<sup>16 18</sup> 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.<sup>21</sup> 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.<sup>9 22</sup> In axSpA also, it appears patients and physicians may value disease activity differently.<sup>10-12</sup> 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.<sup>23</sup> 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.<sup>24</sup>

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.<sup>25</sup> 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.<sup>26</sup>

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.<sup>18</sup> 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).<sup>18–27</sup> 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

<sup>1</sup>Sorbonne Universités, UPMC Univ Paris 06, Institut Pierre Louis d’Épidémiologie et de Santé Publique, GRC-UPMC 08 (EEMOIS), Paris, France

<sup>2</sup>Department of Rheumatology, AP-HP, Pitié Salpêtrière Hospital, Paris, France

<sup>3</sup>Medicine Faculty, Paris Descartes University, Paris, France

<sup>4</sup>Rheumatology B Department, APHP, Cochin Hospital, Paris, France

<sup>5</sup>Department of Clinical Immunology & Rheumatology, Amsterdam Rheumatology Center, Amsterdam and Atrium Medical Center, Heerlen, The Netherlands

<sup>6</sup>Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands

<sup>7</sup>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

- Mengshoel AM. Living with a fluctuating illness of ankylosing spondylitis: a qualitative study. *Arthritis Rheum* 2008;59:1439–44.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Cooksey R, Brophy S, Gravenor MB, et al. Frequency and characteristics of disease flares in ankylosing spondylitis. *Rheumatol Oxf Engl* 2010;49:929–32.
- Brophy S, Calin A. Definition of disease flare in ankylosing spondylitis: the patients’ perspective. *J Rheumatol* 2002;29:954–8.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.