



HAL
open science

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

Laure Gossec, Agnès Portier, Robert Landewé, Adrien Etcheto, Victoria Navarro-Compán, Féline Kroon, Désirée van Der Heijde, Maxime Dougados

► To cite this version:

Laure Gossec, Agnès Portier, Robert Landewé, Adrien Etcheto, Victoria Navarro-Compán, et al.. Preliminary definitions of ‘flare’ in axial spondyloarthritis, based on pain, BASDAI and ASDAS-CRP: an ASAS initiative. *Annals of the Rheumatic Diseases*, 2016, 75 (6), pp.991-996. 10.1136/annrheumdis-2015-208593 . hal-01338693

HAL Id: hal-01338693

<https://hal.sorbonne-universite.fr/hal-01338693>

Submitted on 29 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article for submission to Ann Rheum Dis

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

Laure Gossec, Agnès Portier, Robert Landewé, Adrien Etcheto, Victoria Navarro-Compán, Féline Kroon, Désirée van der Heijde, Maxime Dougados

Authors

Laure Gossec, MD, PhD, Sorbonne Universités, UPMC Univ Paris 06, Institut Pierre Louis d’Epidémiologie et de Santé Publique, GRC-UPMC 08 (EEMOIS); AP-HP, Pitié Salpêtrière Hospital, Department of rheumatology, Paris, France

Agnès Portier, MD, Paris Descartes University, Medicine Faculty; APHP, Rheumatology B Department, Cochin Hospital, Paris France; Sorbonne Universités, UPMC Univ Paris 06, Institut Pierre Louis d’Epidémiologie et de Santé Publique, GRC-UPMC 08 (EEMOIS) Paris France

Robert Landewé, MD, PhD, Department of Clinical Immunology & Rheumatology, Amsterdam Rheumatology Center, Amsterdam and Atrium Medical Center, Heerlen, The Netherlands

Adrien Etcheto, MSc, Paris Descartes University, Medicine Faculty; APHP, Rheumatology B Department, Cochin Hospital, Paris France

Victoria Navarro-Compán, MD, PhD, Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands and Rheumatology Department, University Hospital La Paz, Madrid, Spain.

Féline Kroon, MD, PhD, Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands

Désirée van der Heijde, MD, PhD, Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands

Maxime Dougados, MD, Paris Descartes University, Medicine Faculty; APHP, Rheumatology B Department, Cochin Hospital, Paris France

Key words

Axial spondyloarthritis, disease activity, outcome measures, flare, worsening, ASAS

Grants: supported by ASAS.

Competing interests: None relevant to the present work

Authorship:

Maxime Dougados, Laure Gossec, Désirée van der Heijde and Robert Landewé designed the study and oversaw the vignette exercise.

Agnes Portier performed the systematic literature review with the help of Feline Kroon and Victoria Navarro-Compan.

Agnes Portier, Laure Gossec and Adrien Etcheto performed the vignette exercise.

Adrien Etcheto performed the statistical analyses for the vignette exercise.

All authors participated in drafting the article and gave final approval.

Corresponding author

Pr. Laure Gossec, Hôpital Pitié-Salpêtrière, Service de Rhumatologie, 47-83, boulevard de l'Hôpital - 75013 Paris France

laure.gossec@aphp.fr

Tel: [+33 1 42 17 84 21](tel:+33142178421)

Fax: +33 1 42 17 79 59

Running title: Flare in axSpA

Word count: 3915, 3 tables, 27 references, one online figure and 3 online tables.

Abstract N=250 words

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 axial spondyloarthritis (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 randomized controlled trials in axSpA. (2) Vignette exercise: 140 scenarios were constructed for a typical axSpA patient seen at 2 consecutive visits. Each scenario included a change in one of the following outcomes: pain, Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), BASDAI plus CRP, or ASDAS-CRP. Each ASAS expert determined if every scenario from a random sample of 46 scenarios was considered a flare (yes/no). ROC analyses were applied to derive optimal cutoff 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 cutoffs for pain (N=3), BASDAI (N=5), and ASDAS-CRP (N=4) were chosen, with a range of sensitivity 0.60 to 0.99 and range of specificity 0.40 to 0.94 against the expert's opinions.

Conclusions: This data-driven ASAS consensus process has led to 12 preliminary draft definitions of flare in axial SpA, 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 2 situations; in 'flare-design trials', trial treatment is introduced only in case of flare, being the consequence of interruption of the ongoing/previous treatment (e.g. in axSpA if non steroidal anti inflammatory drugs (NSAIDs) have been stopped);(2) and in tapering or discontinuation trials, if the treatment (e.g. Tumor Necrosis Factor inhibitors, TNFi) 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 harmonizing 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 axial SpA 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 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 is exploring differences in the perception of flares by physicians and patients, with the objective to develop a specific outcome measure, i.e. 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 2 main steps to collect data: a systematic literature review and a case-vignette exercise. This was followed by a consensus step.

Systematic literature review

Data retrieval:

First, to gain an overview of flares, studies specifically focusing on flares in axSpA patients, 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 systematic literature review was performed to collect all the definitions of flare used in randomized controlled trials (RCTs) of NSAIDs or TNFi in axSpA patients, up to May 2014. The search was based on 2 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 www.clinicaltrials.gov were also analyzed. 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 1.

Data selection: one investigator (AP) selected all the studies referring to the concept of flare, in adult axSpA patients.

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 cutoff-level, if flare was measured by a combination of several instruments or as a single instrument only, and if flare was conceptualized 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, cutoff 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 1 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 3 authors (LG, AP, and MD) based on only one scenario. Full information is given in online supplementary Table 3. 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 2 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 2 time points, i.e., a flare is an absolute change between 2 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 axial SpA, range 0-10); (b) Bath Ankylosing Spondylitis Disease Activity Index (BASDAI(15) range 0 to 10); (c) C-reactive protein (CRP) as a continuous result in mg/l, coupled to change in BASDAI; and (d) ASDAS-CRP (the Ankylosing Spondylitis Disease Activity Score – CRP(16)) 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 (e.g. 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 (Table 1 and supplementary online Table 3). An example of a vignette for BASDAI is the following: 'A 32 year-old man with a well-established diagnosis of axial SpA consults you at 2 successive time points. 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.

Initially, variations in CRP alone, as well as in NSAID intake (i.e., 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 2 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 to 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).

Analysis

For each outcome separately, the vignettes were analysed per stratum of change in outcome, i.e., for an absolute change of outcome of at least X (thus all vignettes with a BASDAI increase of at least 3 points were analyzed together, then all vignettes with an increase of at least 4 points, etc). 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 (e.g., 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 curves (ROC-curves) were constructed. Areas under the ROC curve were calculated and optimal cutoff values for defining a flare were established. The corresponding sensitivities, specificities, positive predictive values (PPV) and negative predictive values (NPV) were then calculated. The sensitivity is the proportion with (for example) 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

Systematic literature review of definitions used for flare in axSpA studies

A total of 1,013 articles initially screened resulted in 38 studies using some definition of flare in axSpA (online supplementary Table 2). There were 23 randomized clinical trials (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% (online supplementary Table 2). The 2 most frequent definitions used were: absolute BASDAI $\geq 4/10$ with absolute physician assessment $\geq 4/10$ used in 6 studies, and increase in pain $\geq 30\%$ with absolute pain $\geq 4/10$ used in 6 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 at least 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 cutoff of 2.1 (absolute value).

Five studies (13%) used elements of physical assessment and 4 (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 analyze. 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 supplementary online Figure 1. 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 2 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 2 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 ≥ 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 ≤ 4 . (d) Considering a pain change ≥ 3 , more than 80% of the doctors will consider there is a flare if the final value ≥ 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 2 preliminary definitions based on pain: (a) an increase in pain ≥ 2 AND an observed value at the time of the flare ≥ 4 ; (b) an increase in pain ≥ 3 . The performances of these cut-off values are given 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 2 cut-offs for BASDAI (on a 0-10 scale): increase in BASDAI ≥ 2 points and increase in BASDAI ≥ 3 points. For these 2 cutoffs, 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 lower or equal to 4. (d) Considering a BASDAI change ≥ 3 , more than 70% of the doctors will consider there is a flare if the observed value is greater than or equal to 4 or 5.

Thus the selected preliminary cut-offs for BASDAI are based on an increase of at least 2 or at least 3 points, with or without an observed value of at least 4 (Table 3). An additional (combined) definition was derived during the consensus process as follows: if the observed value of BASDAI is ≥ 4 , 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 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 3 cut-offs for ASDAS-CRP changes: increase in ASDAS-CRP ≥ 0.6 , 0.9 or 1.1. For these 3 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) ≥ 1.3 (i.e., 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, e.g. 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 (e.g. men vs women; or patients with vs those without extra-articular manifestations). 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 axial SpA.(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 axial SpA. However the interpretation of CRP variations alone (i.e., in the absence of concomitant changes in symptoms) was difficult, giving rise to discussions e.g. 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, e.g. remission. This study does not explore the patient's perspective on flares. Ongoing work in rheumatoid arthritis 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 cutoff 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 indications, in the vignette exercise, to the ASAS experts of what they should consider to be a flare (e.g., 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 worsening 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 (i.e. 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 (i.e. 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 rheumatoid arthritis, 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 (i.e., 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 generalizability.

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 2 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.

References

1. Mengshoel AM. Living with a fluctuating illness of ankylosing spondylitis: a qualitative study. *Arthritis Rheum*. 2008 Oct 15;59(10):1439–44.
2. Trijau S, Avouac J, Escalas C, Gossec L, Dougados M. Influence of flare design on symptomatic efficacy of non-steroidal anti-inflammatory drugs in osteoarthritis: a meta-analysis of randomized placebo-controlled trials. *Osteoarthr Cartil OARS Osteoarthr Res Soc*. 2010 Aug;18(8):1012–8.
3. Sieper J, Lenaerts J, Wollenhaupt J, Rudwaleit M, Mazurov VI, Myasoutova L, 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 Jan;73(1):108–13.
4. Cantini F, Niccoli L, Cassarà E, Kaloudi O, Nannini C. 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, 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.
6. Sieper J, Rudwaleit M, Baraliakos X, Brandt J, Braun J, Burgos-Vargas R, et al. The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis*. 2009 Jun;68 Suppl 2:ii1–44.
7. Hermann K-GA, Baraliakos X, van der Heijde DMFM, Jurik A-G, Landewé R, Marzo-Ortega H, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. *Ann Rheum Dis*. 2012 Aug;71(8):1278–88.
8. Bartlett SJ, Bykerk VP, Cooksey R, Choy EH, Alten R, Christensen 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 Feb 15;
9. Bartlett SJ, Hewlett S, Bingham CO, Woodworth TG, Alten R, Pohl C, et al. Identifying core domains to assess flare in rheumatoid arthritis: an OMERACT international patient and provider combined Delphi consensus. *Ann Rheum Dis*. 2012 Nov;71(11):1855–60.
10. Cooksey R, Brophy S, Gravenor MB, Brooks CJ, Burrows CL, Siebert S. Frequency and characteristics of disease flares in ankylosing spondylitis. *Rheumatol Oxf Engl*. 2010 May;49(5):929–32.
11. Brophy S, Calin A. Definition of disease flare in ankylosing spondylitis: the patients' perspective. *J Rheumatol*. 2002 May;29(5):954–8.
12. Stone MA, Pomeroy E, Keat A, Sengupta R, Hickey S, Dieppe P, et al. Assessment of the impact of flares in ankylosing spondylitis disease activity using the Flare Illustration. *Rheumatol Oxf Engl*. 2008 Aug;47(8):1213–8.
13. Godfrin-Valnet M, Puyraveau M, Prati C, Wendling D. Flare in spondyloarthritis: Thresholds of disease activity variations. *Jt Bone Spine Rev Rhum*. 2015 May;82(3):192–5.

14. Veloski J, Tai S, Evans AS, Nash DB. Clinical vignette-based surveys: a tool for assessing physician practice variation. *Am J Med Qual Off J Am Coll Med Qual*. 2005 Jun;20(3):151–7.
15. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. *J Rheumatol*. 1994 Dec;21(12):2286–91.
16. Lukas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. *Ann Rheum Dis*. 2009 Jan;68(1):18–24.
17. Machado P, Navarro-Compán V, Landewé R, van Gaalen FA, Roux C, van der Heijde D. 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 Hoboken NJ*. 2015 Feb;67(2):408–13.
18. Machado P, Landewé R, Lie E, Kvien TK, Braun J, Baker D, et al. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. *Ann Rheum Dis*. 2011 Jan;70(1):47–53.
19. Banks J, Hollinghurst S, Bigwood L, Peters TJ, Walter FM, Hamilton W. Preferences for cancer investigation: a vignette-based study of primary-care attendees. *Lancet Oncol*. 2014 Feb;15(2):232–40.
20. Das J, Kwan A, Daniels B, Satyanarayana S, Subbaraman R, Bergkvist S, et al. Use of standardised patients to assess quality of tuberculosis care: a pilot, cross-sectional study. *Lancet Infect Dis*. 2015 Aug 7;
21. Dougados M, Simon P, Braun J, Burgos-Vargas R, Maksymowych WP, Sieper J, et al. ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. *Ann Rheum Dis*. 2011 Feb;70(2):249–51.
22. Hewlett S, Sanderson T, May J, Alten R, Bingham CO, Cross M, 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 Oxf Engl*. 2012 Jan;51(1):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 Jan;32(1):80–5.
24. Tubach F, Wells GA, Ravaud P, Dougados M. Minimal clinically important difference, low disease activity state, and patient acceptable symptom state: methodological issues. *J Rheumatol*. 2005 Oct;32(10):2025–9.
25. Curtis JR, Yang S, Chen L, Pope JE, Keystone EC, Haraoui B, et al. Determining the minimally important difference in the clinical disease activity index for improvement and worsening in early rheumatoid arthritis. *Arthritis Care Res*. 2015 May 18;
26. van der Heijde D, Sieper J, Maksymowych WP, Dougados M, Burgos-Vargas R, Landewé R, 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 Jun;70(6):905–8.

27. Wariaghli G, Allali F, Idrissi Z, Berrada K, Hmamouchi I, Abouqal R, et al. Evaluation and stability of the Patient Acceptable Symptom State (PASS) over time in patients with ankylosing spondylitis. *Clin Exp Rheumatol.* 2012 Feb;30(1):106–9.

Table 1. The outcome changes used in the vignette exercise

Outcomes	BASDAI (0-10)	Pain due to axial SpA (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 to 6	0 to 6	0.6 to 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 to 20 mg/l
Total number of vignettes: 140	49	49	24	18

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 response	1 (2.6%)	Alone	Loss (N=1)
Isolated patient-reported outcome	Pain (0-100mm)	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 $\geq 4/10$ (N=1) Rel. change $\geq 2/10$ (N=1)
	NSAID intake	1 (2.6%)	Alone	Presence
Physician assessment or	Physician global	7(18.4%)	Combination N=7	Absolute value ≥ 4

laboratory value	assessment (0-10)			(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)

BASDAI: Bath Ankylosing Spondylitis Disease Activity Index (14); ASDAS: Ankylosing Spondylitis Disease Activity Score(16); ASAS: Assessment of SpondyloArthritis; NSAIDs: Non steroidal Anti Inflammatory Drugs; Rel: relative; Abs: absolute.

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

AUC: area under the ROC curve; Δ : change; Se: sensitivity; Spe: specificity; PPV and NPV: positive and negative predictive values.