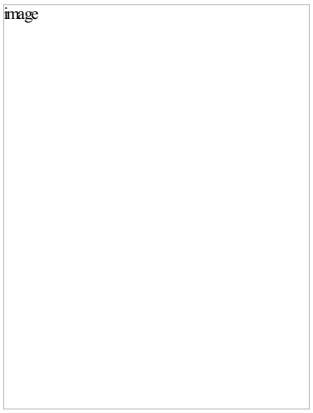


How is symptom flare defined in musculoskeletal conditions: A systematic review

Nathalia K. Costa, Manuela L. Ferreira, Marita Cross, Joanna Makovey, Paul W. Hodges

www.elsevier.com/locate/bios

PII: S0049-0172(17)30570-X
DOI: https://doi.org/10.1016/j.semarthrit.2018.01.012
Reference: YSARH51303

To appear in: *Seminars in Arthritis and Rheumatism*

Cite this article as: Nathalia K. Costa, Manuela L. Ferreira, Marita Cross, Joanna Makovey and Paul W. Hodges, How is symptom flare defined in musculoskeletal conditions: A systematic review, *Seminars in Arthritis and Rheumatism,* doi:10.1016/j.semarthrit.2018.01.012

Title: How is symptom flare defined in musculoskeletal conditions: A systematic review

Authors: Nathalia K. Costa, BPhty[a]; Manuela L. Ferreira, PhD[b]; Marita Cross, PhD[b]; Joanna Makovey, PhD[b], Paul W. Hodges, PhD[a]

Affiliations: [a] The University of Queensland, School of Health and Rehabilitation Sciences, Brisbane, QLD, Australia; [b] The University of Sydney, Institute of Bone and Joint Research, The Kolling Institute, Sydney Medical School

Source of support: This study is supported by a Program grant (APP1091302); Centre of Research Excellence grant (APP1079078) and Fellowship (PH – APP1102905) from the National Health and Medical Research Council (NHMRC) of Australia. MLF is supported by the Sydney Medical School Foundation Fellowship, Sydney Medical School/The University of Sydney.

Disclaimers: The authors have no conflicts of interest to declare.

Corresponding author:

Dr Paul Hodges

Centre of Clinical Research Excellence in Spinal Pain, Injury and Health, School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Queensland, 4072 Australia

E-mail: p.hodges@uq.edu.au

Tel: +61 404 854 589

Abstract

Objective: To systematically review the definitions for "flare" in musculoskeletal conditions, the derivation processes, and validation of definitions for the 12 most burdensome musculoskeletal conditions.

Methods: A literature search was conducted in MEDLINE, EMBASE, CINAHL, AMED, PsycInfo and Lilacs to identify studies that investigated derivation or validation of a flare definition, which we considered as a phrase or group of domains.

Results: Reports of derivation of flare definitions were identified for 9/12 musculoskeletal conditions. Validation of flare definitions was initiated for 4/12. For each condition, different derivation and validation methods have been used, with variable levels of consumer involvement, and in some cases different groups have worked on the process in parallel. Although some flare definitions began simply as "symptom worsening" or "change in treatment", most evolved into multidimensional definitions that include: pain, impact on function, joint symptoms, and emotional elements. Frequently initial attempts to create phrase to define the term flare evolved into consensus on the breadth of domains involved. Validation has compared flare definitions/domains against measures of disease activity, clinicians' diagnosis, response to drug therapy, or a combination.

Conclusion: This review suggests that greater characterisation and definition of flares in musculoskeletal conditions are linked to the inclusion of multiple perspectives, multifaceted domains and compound comparators for their validation. Further work is required to optimise and test the derived definitions for most musculoskeletal conditions. As some elements are disease-specific flare definitions cannot be extrapolated to other conditions. Research regarding flare in back pain (most burdensome disease) is limited.

Keywords: Systematic review, flare, definition, musculoskeletal diseases.

1. Introduction

Musculoskeletal conditions are pervasive and a leading contributor to the global burden of disease (1). As many musculoskeletal conditions do not have a cure and become lifelong problems, research has focussed on identification of factors that influence the progression from acute to chronic conditions, the determinants of the rate of progression of the disease, and possible clinical or environmental interventions to halt the progression or to reduce the impact of disease. Such consideration depends on a clear understanding of the time-course of disease and this relies on clear and unambiguous measures of disease state. A major issue is that although symptoms are ongoing, most are characterised by variation or fluctuation of symptoms (2) but not all fluctuations are likely to be important (3, 4). For several conditions, periods of increased severity of the condition are referred to as a "flare" or "flare-up". Despite the frequent use of this term in research and clinical practice, it is rarely clearly defined. It is difficult to be certain whether a flare has the same meaning for different conditions. Furthermore, it is unclear whether the terms are used consistently between patients, clinicians and researchers, or within these groups.

For some conditions, such as rheumatoid arthritis (RA), a detailed process has been initiated to define and understand flare (5, 6). In these cases, there is advanced understanding of patient and clinician interpretations of the term, but the results of each step in the process are published in individual papers (e.g. (7, 8)) making it difficult to clarify the overall derivation process. For other conditions, parts of the process of derivation of a definition has been undertaken, but without overarching coordination of the process. Consequently, the term flare is used for multiple purposes such as an outcome measure for clinical trials without an explicit definition in most cases. To fully ascertain the current understanding of flare and its use in research and clinical practice it was necessary to undertake a systematic review of the literature.

The overall aim of this systematic review was to comprehensively review the definitions that have been derived and validated for the term flare (or flare up) in the 12 most burdensome musculoskeletal conditions defined in the Global Burden of Disease Study (9, 10). Our interest was to gain a comprehensive understanding of flare definitions and domains used to define a flare (i.e. worsening of condition; change in treatment). Our specific aims were to: (i) document and contrast the definitions or domains used to identify and/or characterise flare for the most burdensome musculoskeletal conditions; (ii) assess the methods used to derive the definitions or domains; and (iii) review studies that assess the validity of definitions or domains of flare.

2. Methods

2.1 Search strategy

The methods of this review have been registered with PROSPERO (CRD42017056996). We performed a systematic search to identify studies of the key musculoskeletal conditions that derived or validated a definition for flare. The key musculoskeletal conditions considered were the "major" and "other musculoskeletal conditions" determined by Global Burden of Disease (GBD) 2010 Study (10, 11). The "major musculoskeletal conditions" are the ones more likely to contribute to the largest proportion of musculoskeletal burden based on daily-adjusted life year, i.e., hip and/or knee osteoarthritis (OA), low back pain (LBP), rheumatoid arthritis (RA), gout and neck pain. The group of "other musculoskeletal disorders" includes systemic lupus erythematosus (SLE), ankylosing spondylitis (AS), psoriatic arthritis (PsA), juvenile rheumatoid arthritis (jRA), osteomyelitis, fibromyalgia and shoulder pain. We used the search term "flare" combined with each of those conditions. A systematic search was conducted in MEDLINE, EMBASE CINAHL, AMED,

PsycInfo and Lilacs from the earliest record to February 2017. Searches were conducted in 2 phases; December 2015 and February 2017. The search terms were adapted for use according to database-specific filters. No restriction was applied on study design or language.

2.2 Study selection

The results were exported into an EndNote X7.0 database, and duplicates were removed. Studies on animals, pregnant women, participants with non-musculoskeletal conditions (e.g. cancer) or undergoing surgery were excluded. Two independent reviewers (NC and MF) screened all titles and abstracts for potential inclusion. Disagreement was resolved by discussion. Full texts of all potentially eligible studies were evaluated for inclusion by two reviewers (NC and PH) on the probability that (1) the study defined flare, or (2) attempted to validate flare definitions or domains. To capture studies which defined flare, we considered the term "defined" to broadly mean both a "phrase" that provided a formal definition of the state of flare or a multi-dimensional tool/group of "domains" that could be considered to determine whether or not a person was experiencing a flare. Studies including mixed patient populations (e.g. non-musculoskeletal conditions and musculoskeletal conditions) and patient populations with non-musculoskeletal symptoms (e.g. skin features in psoriatic arthritis) were only included if the flare data were separately presented for musculoskeletal conditions. Studies written in languages other than English were eligible if a translator was available. However, no eligible study was found in non-English languages. Studies including induced flares were excluded.

2.3 Study grouping

Identified studies were grouped according to the following purposes:

GROUP A: Studies that described an experimental study or series of studies that aimed to *derive* a definition or set of domains that characterise a flare;

GROUP B: Studies that aimed test the *validity* of a definition that had been proposed/derived for flare or a set of domains that characterise flare.

2.4 Assessment of study methodological quality

Study methodological quality was assessed as it pertains to the purpose of our review (derivation and validation of the definition for flare) and not the principal design of the study (e.g. if data from an RCT was included, the design of the RCT was not evaluated, but the process related to derivation of the definition of flare was considered). Different aspects of study quality were considered for each study group.

A major consideration for papers in Group A was that for most conditions, the entire process used for derivation of a definition was not contained within a single paper, but instead described in a series of papers. The methodological quality was appraised in terms of whether the process:

- Provided clear identification of disease or condition;
- Considered the perspective of patients, clinicians, and/or researchers (groups were recorded);
- Involved an experimental method to derive the definition (recorded as Delphi; qualitative study; consensus meeting; etc.);
- Considered single or multiple domains to define flare (domains were recorded);
- Used an experimental method to identify a threshold/cut-off score on a symptom scale to characterise a flare (if a cut-off was used this was recorded as “yes” or “arbitrary” cut-off);
- Involved a method to reach consensus for the definition/domains (recorded as yes/no).

Methodological quality of studies in Group B was appraised using the following criteria:

- Clear identification of disease or condition;
- Representativeness of the sample (i.e. sampled from general community, or sampled from a specific group of patients);
- Clear description of comparator (comparator recorded);
- Blinding of patients and investigators to assessment of comparators used to validate a proposed definition (recorded as yes/no/not applicable);
- Inclusion of follow-up long enough to allow for flares to be experienced (recorded as yes/no/not applicable).

2. 5 Data extraction

The data extracted from the studies differed according to group. For all studies we recorded the authors, year of publication and condition studied. For studies in Group A we recorded the methods used to derive and/or definition or domains, the contributing groups (i.e. consumers, clinicians), whether consensus for the definition/domains was achieved, the flare definition or identified domains, and measurement tools when appropriate. All definitions or groups of domains used to characterise flares included reference to multiple features, and the terms used to describe these differed between diseases and research groups. To assist with comparison of the definitions/domains, we undertook a thematic analysis and allocated each key elements or features of the definitions to one of eleven themes identified by consensus of the author group (Table 1).

For studies in Group B we documented the definition and/or domains of a flare, the methods used to test its validity, the groups involved in the validation process, whether the definition/domains were validated, and the comparator used for the validation of the definition. Data extraction was conducted by independently by reviewers (NC, PH and MF) and in case of disagreement, consensus was reached by discussion.

2.6 Data presentation

Data on derivation and validation processes are presented separately for each musculoskeletal condition in Table 2 and 3. Study details and assessment of study quality for Groups A and B are presented separately in Tables 4 and 5, respectively.

3. Results

Figure 1 presents the flow chart of the papers screened and included in this review. Data are presented according to the two a priori defined categories below.

3.1 GROUP A: Derivation of a definition or domains of flare.

Twenty-eight studies described the derivation of a definition of flare or described the identification of domains that characterise a flare. Most considered rheumatologic conditions (RA – 9; jRA – 1; Gout – 3; SLE – 6; Juvenile SLE – 1; PsA – 2; OA – 1; AS – 3; Fibromyalgia - 1) and one study derived a definition for flare in LBP. Diverse methods have been used to develop definitions. All studies clearly identified the condition and included an experimental method in their derivation, and the derived definition included at least two domains. Forty-six per cent of studies included perspectives from more than one interest group. Forty-three per cent involved a method to reach consensus for the definition/domains. Only 21% of studies used an experimental method to identify the cut-off on a symptom scale. Figure 2 shows the number of conditions that include consideration of each of the themes in the definition. The themes of the most commonly included domains were “pain”, “impact on function” and “joint symptoms”. Figure 3 shows the number of themes considered for each condition. RA, SLE and AS consider the broadest range of themes in their definition. The process undertaken, the methodological quality and the resulting definition or domains differed between conditions and are presented below separately by condition.

Rheumatoid Arthritis (RA)

The Outcome Measures in Rheumatology (OMERACT) initiative has coordinated the derivation of a definition of flare in RA, published across nine papers and one abstract (four papers considered validation, see Group B). An initial working definition was derived at the 9th OMERACT meeting in 2008 (OMERACT9) and focused on the domains "worsening of condition", "duration", "symptom intensity", "change in treatment" (5-7). Analyses of focus groups of patients in 5 countries showed that global visual analogue scale (VAS) and joint count do not adequately capture flares, suggesting a need for a deeper understanding of this experience (12). More recent publications describe a change in strategy from a phrase to define flare, to the identification of multiple domains of disease activity (8) that form the foundation for a questionnaire to indicate the presence of a flare. The Preliminary Flare Questionnaire (PFQ) was developed in 2012 at OMERACT11 (13) and using core flare domains identified in previous meetings. There was lack of consensus between views of consumers and clinicians regarding flare domains (e.g. patients, but not, clinicians considered fatigue to be important). Validation of the Rheumatoid Arthritis Flare Core Domain Set and the Preliminary Flare Questionnaire (PFQ) are discussed in Group B.

In parallel, the Strategy of Treatment in Patients with Rheumatoid Arthritis group of the French Society of Rheumatology derived a definition of RA flare through the development of a self-administered questionnaire (FLARE-RA) containing multiple partially distinct domains (14, 15). The FLARE-RA has been validated (see Group B) and a threshold of 2.5 is suggested for identification of RA flare (16).

Juvenile Rheumatoid Arthritis (JRA)

Brunner et al. (17) used data from a randomised controlled trial of a disease-modifying drug (Etanercept) to propose and assess candidate flare definitions. Participants in the placebo group were assumed to experience flares during the trial duration, whereas those in the treatment arm were assumed to be free from flare. The combination of scores from a group of clinical variables (core response variables; CRV) was tested using receiver operator characteristic (ROC) and positive and negative predictive values for group allocation. Three proposed definitions performed adequately for accurate identification of presence of a flare. These related to the domain of "worsening of condition" as assessed by clinical tests, in combination with results from questionnaires and laboratory tests.

Gout

Although flare is frequently identified as an important feature of chronic gout (18-20) there have be few attempts to define or identify the domains that characterise a gout flare. A consensus exercise involving two Delphi surveys and a cognitive mapping process involving patients, clinicians and experts identified nine key elements of a gout flare definition (21). The European League Against Rheumatism (EULAR) evaluated the utility of these elements of gout flare to develop a patient-centred definition. At eight international sites, 210 patients were assessed by nine experienced rheumatologists. The physician's determination of the presence or absence of gout flare was the comparator. Logistic regression and classification and regression tree (CART) were used to identify the best predictors of a flare. Patient-reported flare, presence of any warm joints or swollen joints, and joint pain at rest with a score >3 were strongly associated with the comparator physician's determination of presence of a gout flare (22).

A novel approach to flare definition was attempted by MacFarlane et al (23). Using computerised medical records, they aimed to identify claims-based algorithms that could to identify gout flares and assessed against physician documentation of gout flare. Claims-based algorithms did not accurately identify gout flares.

Systemic Lupus Erythematosus (SLE)

Flares have been largely recognised as an important feature of the disease pattern in SLE and used to represent "worsening of disease activity" (24). Eight instruments in different versions and modifications (e.g. British Isles Lupus Assessment Group 2004 index – BILAG, Systemic Lupus Erythematosus Disease Activity Index – SLEDAI, The Safety of Estrogen in Lupus Erythematosus National Assessment – SELENA, Lupus Activity Index - LAI) have been proposed. As SLE involves many systems other than the musculoskeletal system, most definitions include a multi-system perspective (e.g. seizure, cardiorespiratory symptoms, etc). Most research has used a change in score on these disease activity instruments to quantify and characterize flare (24). Cut-offs for a flare were established for BILAG, SLEDAI, SELENA SLEDAI and LAI (25-29).

In 2006, The Lupus Foundation of America (LFA) convened an International Consensus Panel "Definition and Validation of Lupus Flares" to evaluate needs in defining and measuring SLE flares. Two web-based Delphi surveys of physicians, a second consensus conference (which included patients), and a third Delphi survey were undertaken to reach final consensus (30). The definition of flare was broadened to include more domains as a result of this process. A separate study achieved a consensus for a flare definition and candidate flare criteria in Juvenile Systemic Lupus Erythematosus (jSLE) (31). Work to assess validity of definitions and instruments has been done (see Group B).

Psoriatic arthritis (PsA)

Moverley et al. (32) aimed to define flare in PsA. Interview data from patients were analysed thematically. Flares were defined as a collection of interacting "physical", "psychological" and "emotional" symptoms defined by several domains. Furthering this work, the Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA) conducted two Delphi surveys and a face-to-face discussion to define flare and its domains towards the development of a questionnaire to identify flare in PsA (33). Development of the questionnaire is ongoing.

Osteoarthritis (OA)

Marty et al. (34) conducted two observational cross-sectional studies to derive and validate a diagnostic score for OA flare based on clinical and radiological features. Initial data were obtained from general practitioners and used to develop a diagnostic score for patients seeking primary care. To build the instrument and derive cut-off scores, patients with stable condition and those experiencing a flare were recruited. Patients' demographics and clinical characteristics, including pain severity, functional impairment and treatment, were included in logistic regression models to identify factors independently associated with flare. A score ≥7 was identified when each factor was given a weight for the odds of not having a flare. This score was validated using a rheumatologist database (see Group B).

Ankylosing Spondylitis (AS)

Brophy et al. (35) examined patients' perceptions of the important domains for flare definition in AS. In twenty group meetings of 7-12 participants, 214 patients were asked questions regarding flare and its triggers, duration, sequelae and frequency. A consensus definition of flare was obtained by each group. Data revealed two forms of flares ("localised

to one area" and "throughout the body"). Both were marked by highly painful symptoms triggered by stress in the majority of cases.

The Assessment of Spondyloarthritis (ASAS) group conducted a systematic review of flare definitions and vignette exercises to develop a consensus definition of flare. Diverse flare definitions were found and 12 preliminary draft flare definitions were proposed for further validation (36).

Godfrin-Valnet et al. (37) aimed to derive a definition of flare based on disease activity indices. ROC curves were used to identify thresholds for optimum sensitivity and specificity to detect flares reported by patients and physicians. Specificity was strong, but sensitivity was only moderate and agreement between patients and physicians was only moderate (Kappa 0.68).

Fibromyalgia

Vincent et al. (38) identified domains of flares in patients with fibromyalgia using a qualitative survey. Content analysis identified three main content areas and key themes within each of these content areas: causes of flares (stress, overdoing, poor sleep and weather changes), flare symptoms (flu-like symptoms, pain, fatigue, other symptoms) and dealing with flares (treatments, rest, avoid everything and wait it out). No definition was derived or validation attempted.

Low Back Pain

One study has proposed a definition of flare for LBP (39). This qualitative study originally aimed to examine how individuals with a history of LBP describe events that could be conceptualized as "recurrent episodes". The results suggested people with LBP believe they live with a baseline level of their condition, and in their opinion, a recurrent episode is better represented as a period worsening of their ongoing condition, recognised as "flare-ups"

or “attacks”. According to participants’ views, flares were not only characterised by pain, but also moderate activity limitations, participation restrictions, need for activities to help manage their pain, and fear of worsening of their condition. “Flares” tended to be viewed as manageable whereas “attacks” were considered more severe, involving more pain and “the inability to do anything”.

3.2 GROUP B: Validation of a flare definition

Twenty-three studies undertook a process to validate the definition or domains of flare as summarised in Table 5. All considered rheumatologic conditions (RA – 14; Gout – 1; SLE – 6; OA – 2). Only 56% of studies recruited participants from the general community. Others used specific groups (30%; e.g. participants receiving a specific treatment) or did not provide this information (14%). Of studies aiming to validate a flare definition, 65% (n= 15) were observational studies and 21% (n=5) were randomised clinical trials. One study was a multicentre trial and one study aimed to validate the flare definition in different languages. Disease activity measures were used as the comparator in 65% (n=15) of studies. From those, 27% (n= 4) used disease activity measures combined with other instruments to measure function, health-related quality of life and/or outcome measures of other domains, and 53% (n= 8) combined disease activity measures with either patients’ and/or clinicians’ perspectives. Eighteen per cent (n=4) relied solely on clinicians’ perspectives, and 9% (n=2) considered only patients’ perspectives. In 73% of studies patients and/or investigators were blinded to results of the comparator in the validation process. Results are presented below separately by condition.

Rheumatoid Arthritis

Responsiveness of the OMERACT preliminary flare questions (PFQ) to changes in flare status determined by rheumatologist, patient self-reported state or one of the two proposed DAS28-based criteria was tested. Scores in each PFQ domain changed significantly for those experiencing flares and remained unchanged for whose flare status was unchanged. DAS28 criteria (Ritchie Articular Index, number of swollen joints, erythrocyte sedimentation rate (ESR) and a general health assessment scored on a VAS (40)) were insensitive to change (41). Construct and discriminatory validity of the PFQ were examined in a RCT of TNF-blocking drugs vs. usual care. A statistically significant change in scores (compared to baseline) of all PFQ items was found only in patients experiencing a flare defined by DAS28 (42). Breakout group discussions at OMERACT 2014 identified issues regarding assessment of the domains stiffness and self-management. Refinement of RA flare measurement continues (43). Convergent and divergent construct validity of the individual domains of flare identified by OMERACT has been assessed (44) from their change over time (3 and 6 months) and compared against flare defined by DAS28 and three flare working definitions: *worsening of symptoms (self-reported), change in treatment (observed)* and the combination of these two domains. Flare defined as 'self-reported worsening of symptoms' yielded higher standardised mean differences than "change in treatment".

Barlett et al. showed discriminant validity of PFQ between patients with and with self-reported flare, and convergent validity between PFQ and validated RA measures (45). Bykerk and colleagues showed modest agreement between patients and rheumatologists in identifying a flare and the concordance of clinical and patient-reported outcomes (PRO) with flare status; patients identifying more swollen and tender joints. PRO significantly discriminated between patients reporting flare and those who did not (46)

Validity of the FLARE instrument has been assessed in six studies (47-52). First, a prospective trial of the FLARE Self-report questionnaire showed it is a valid and valuable instrument to detect RA flare between visits to the physician (50). Second, an observational study showed good correlation between detection of past flares using the FLARE-RA questionnaire and the Routine Assessment of Patient Index Data (RAPID) questionnaire (47). Third, comparison of the questionnaire data against clinical and laboratory measures of disease activity and severity revealed a high correlation with the all measures except for the Physician Clinical Assessment (PCA) (48). Fourth, validity and reliability of the Danish version of the Flare Instrument (FI) was found to be excellent against DAS28 (intra class correlation coefficient >0.95) (52). Fifth, assessment of the criterion and concurrent validity of the Danish version of the FI against DAS28-CRP found good to moderate diagnostic properties (49). Sixth, FI was translated into Spanish and a cut-off value ≤50.5 was found to determine the presence of a disease flare with high sensitivity and specificity (>60), which was more strongly correlated with patient's opinion and treatment change between visits than with physician's opinion and disease activity scores (51).

Two studies validated the definition of flare based on DAS28. Dougados et al. used RCT data to compare DAS28 against the investigators' opinion. DAS 28 was sensitive (88-100%) but not specific (57-65%) to identification of a flare (53). Portier et al. (54) showed difference in conceptualisation of flare between patients and DAS28 (which was considered to reflect the physician's opinion; Kappa 0.44). Patient-reported flares emphasised physical and mental domains, but joint pain was rarely indicated as the only flare symptom (16.8% of the total flares).

Gout

Teoh et al. (55) assessed the validity of six methods to report flare against gout disease activity measures [patient and physician global assessments, joint counts and C-reactive protein (CRP)]. All methods correlated well with measures of gout disease activity, except "time to first flare".

Systemic Lupus Erythematosus (SLE)

Patients with SLE were included in a study examining the intra- and inter-rater reliability of the Physician Global Assessment (PGA), the SLEDAI and SELENA. Whereas the PGA and the SLEDAI showed high intra- (>0.87) and inter-rater (>0.75) reliability, the SELENA tool showed only moderate reliability (<0.52). Agreement among the three instruments was poor (56).

Data from five visits of 230 SLE patients were assessed to determine correlations between SLEDAI scores and clinicians' views of disease activity levels. A flare was considered to have occurred if new or increased therapy was prescribed for active disease, or if the physician's notes indicated an expression of concern or use of the term flare. An increase on SLEDAI over 3 points agreed with the clinical determination of flare (26).

Isenberg et al. (57) compared flare definitions derived from the BILAG 2004 and the SFI against the PGA based on the Lupus Foundation of America (LFA) flare definition. The BILAG 2004 based definition of flares was more consistently associated with the opinions of physicians. Petri et al. (58) compared the revised SELENA flare index (rSFI) and its original version (SFI) to the BILAG index, which discriminates between severe, moderate and mild flares in patients with SLE. The results suggest higher agreement between the BILAG and the rSFI than its original version (SFI). The SELENA-SLEDAI flare index (cSFI) was compared to a version *without a criteria related to medication* and with clinical judgment. Results

indicate that the addition of medication dose did not improve the relationship between cSFI and clinically defined flare (59). Brunner et al. (31) identified patients with flare or stable jSLE using their candidate flare criteria. PGA, disease activity measure, anti-dsDNA, creatinine ratio and parental global assessment of well-being were found to adequately capture jSLE global flares diagnosed by rheumatologists.

Osteoarthritis

Using a qualitative methodology Murphy et al. (60) showed that the definition of osteoarthritis flare based on "inadequate pain relief brought on by too much activity" did not match patients' perspectives and opinions. Marty et al. (34) revealed good validity for Knee Osteoarthritis Flare-Ups Score (KOFUS) and cut-off score to define flare when using the rheumatologist diagnosis as a standard.

4. Discussion

This systematic review identified that a process towards a flare definition has been initiated for nine of the 12 most burdensome musculoskeletal conditions. Diverse methods have been used to derive and/or validate a definition for flare or its domains in musculoskeletal conditions with varying degrees of patient consultation. A process to validate the definitions/domains has been initiated for only four. For some burdensome diseases (e.g. shoulder pain, neck pain and osteomyelitis), no process of definition derivation has been initiated. Research in the field of LBP – the most burdensome disease globally in terms of years lived with disability (9) – is limited to one qualitative study that indicates people with LBP consider their condition to be ongoing and characterised by "flares" rather than discrete episodes of pain, even if they have pain-free periods.

4.1 Themes in flare definitions

Although several processes to define flare began with an objective to derive a simple definition for flare there is consensus that definition requires consideration of multiple domains. Pain, impact on function and joint symptoms were the most common themes (included in 77% of the definitions). The themes emotional symptoms (55%) and fatigue (55%) were frequently considered. Some definitions were uni-dimensional and focused on the biomedical features of objective clinical measures (e.g. (5-7)) whereas others were multidimensional and considered a biopsychosocial profile with inclusion of objectively measured and self-reported aspects of the experience of a flare (e.g. (8, 13)). The methods used and groups included in the derivation processes had an impact on the flare definitions. As might be expected, when clinicians' perspectives were emphasised, clinical signs (e.g. *Gout* (19) - pain and warm, swollen and stiff joints) were highlighted. Definitions derived using clinicians' perspectives also tended to exclude "fatigue" as an important component of flare, but included "change in treatment" or "pain" as main components (e.g. *RA* (13, 44)). In contrast, consideration of patient perspectives revealed flare as a complex experience and the domain "change in treatment" did not adequately identify flares (42, 44, 59). Further, when patients' views were considered, the definitions tended to include a broader biopsychosocial conceptualisation of flare, including features such as changes in function, emotional aspects and fatigue (e.g. RA, SLE, fibromyalgia and osteoarthritis patients (13, 30, 38, 60)). Not surprisingly there is only moderate agreement between patients and physicians regarding when they have a flare (37, 41, 54). Regardless of the group considered, there is consensus that a definition for flare must consider more than an increase in pain. HCPs', researchers' and patients' perspectives appear to be complementary and essential for a deeper understanding of flares. This finding agrees with current literature which argues that the use of patient reported outcome measures narrows the gap between clinicians' and patients'

views of health states, leading to better communication and decision making (61). Automated methods to identify flare from patient records have not been successful (23, 55)

4.2 Validation of flare definitions

Using a variety of approaches to assess the validity of the proposed definitions/domains, most studies showed good relationship with other measures/features of flare. Studies relied mainly on comparison against existing clinical measures of disease activity and only a few include measures of non-biomedical domains (e.g. fatigue, health related quality of life). Thus, in most cases the flare definitions/domains under investigation were more multidimensional than the comparators used to validate them. This questions whether the studies provide an optimal estimate of validity. Despite this limitation, available data show that flare definitions/domains agree with disease activity measures, and correspond to patient's and physician's perspectives, with some exceptions (48, 60)). An issue to consider is that many validation studies relied on participants that are unlikely to represent the general population (e.g. limited to a subset of patients receiving a specific intervention).

4.3 Comparison between conditions

Among musculoskeletal conditions included in this review, the RA OMERACT process towards a flare definition has been the most comprehensive with in-depth consideration of all stakeholders' opinions. Patients have been included in all steps towards a flare definition, including perspectives of patients from international sites. Recognising flare as a multidimensional construct, the RA OMERACT group were the first to change their strategy from derivation of a single phrase to describe the state, to an approach that aims to identify the multiple domains that might characterise a flare (6). The established flare domains are being used to build a questionnaire to identify flare based on defined thresholds (41).

No domain was universally used to define a flare among the musculoskeletal conditions investigated. The heterogeneity among musculoskeletal conditions, the use of different methods and different degrees of maturity of the derivation process to define flare makes it difficult to directly compare definitions applied to different conditions. However, some differences between flare definitions/domains are clear. First, the complex pathology of SLE which affects multiple systems in addition to musculoskeletal issues is reflected in the high number and nature of the domains involved in its flare definition. Second, several diseases have hallmark signs that are included – e.g. number of swollen/painful joints in RA and Gout. Others are less easily identified through clinical signs and do not contain disease-specific domains - e.g. Fibromyalgia and LBP.

It might be reasonable to expect that flare in an ongoing systemic condition such as RA could be different to conditions commonly characterised by repeated episodes such as LBP. However, this was not the case. Contrary to the expectation of Young et al. (39), patients with LBP considered their condition to be ongoing (despite periods without pain) and characterised by flares rather than discrete symptom episodes.

Flare in some of the major musculoskeletal conditions have not yet been considered. As flare is not considered identically for all conditions, it is not possible to extrapolate from existing definitions to conditions where the process has not been commenced and the process must be undertaken individually if flare is to be considered an accurate and relevant aspect of a condition. It is likely that similar to the other investigated musculoskeletal conditions, flares of neck pain, shoulder pain and osteomyelitis may be a multidimensional concept, and aspects of the disease and its impact, other than pain, should be considered.

4.4 Flare vs. other types of symptom fluctuation

This systematic review suggests the concept of flare differs from other definitions commonly used to describe fluctuations of symptoms (e.g. episode, recurrence), as those are commonly defined only taking into consideration duration, intensity of symptoms, and when the period of symptoms is preceded by a period without symptoms (e.g. LBP recurrence (62)). It is likely that risk factors for flares will be different to those associated with a new episode of a condition, and these need to be studied separately.

4.5 Conclusion

The findings of this systematic review indicate that flare is a multifaceted experience that differs in some respects between conditions. There is consensus that definitions of flare require consideration of aspects in addition to pain and symptom intensity, but the breadth of features of the experience of flare was greater when perspectives of patients were considered. Whether flare can be adequately defined using a phrase or requires detailed consideration of multiple domains has not been established and may vary between conditions. It is clear that flare cannot be distilled to a simple consideration of whether treatment has changed. Validation of definitions remains challenging as most measures used for comparison are more restricted in their consideration of flare that the definitions and domains that have been derived. Further work is required to optimise and test the derived definitions for most musculoskeletal conditions and work must commence for some.

References

1. Collaborators GBoDS. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet (London, England). 2015;386(9995):743-800.

2. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003;81(9):646-56.

3. 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. Rheumatology. 2008;47(8):1213-8.

4. Kett C, Flint J, Openshaw M, Raza K, Kumar K. Self-management strategies used during flares of rheumatoid arthritis in an ethnically diverse population. Musculoskeletal Care. 2010;8(4):204-14.

5. Bingham CO, 3rd, Pohl C, Woodworth TG, Hewlett SE, May JE, Rahman MU, et al. Developing a standardized definition for disease "flare" in rheumatoid arthritis (OMERACT 9 Special Interest Group). J Rheumatol. 2009;36(10):2335-41.

6. Bingham CO, 3rd, Alten R, Bartlett SJ, Bykerk VP, Brooks PM, Choy E, et al. Identifying preliminary domains to detect and measure rheumatoid arthritis flares: report of the OMERACT 10 RA Flare Workshop. J Rheumatol. 2011;38(8):1751-8.

7. Alten R, Pohl C, Choy EH, Christensen R, Furst DE, Hewlett SE, et al. Developing a construct to evaluate flares in rheumatoid arthritis: a conceptual report of the OMERACT RA Flare Definition Working Group. J Rheumatol. 2011;38(8):1745-50.

8. Bartlett SJ, Hewlett S, Bingham ICO, 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;71(11):1855-60.

9. Hoy D, March L, Brooks P, Blyth F, Woolf A, Bain C, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis. 2014;73(6):968-74.

10. Smith E, Hoy DG, Cross M, Vos T, Naghavi M, Buchbinder R, et al. The global burden of other musculoskeletal disorders: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis. 2014;73(8):1462-9.

11. Hoy DG, Smith E, Cross M, Sanchez-Riera L, Buchbinder R, Blyth FM, et al. The global burden of musculoskeletal conditions for 2010: an overview of methods. Annals of the Rheumatic Diseases. 2014;73(6):982-9.

12. Hewlett S, Sanderson T, May J, Alten R, Bingham CO, 3rd, 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. Rheumatology. 2012;51(1):69-76.

13. Bykerk VP, Lie E, Bartlett SJ, Alten R, Boonen A, Christensen R, et al. Establishing a core domain set to measure rheumatoid arthritis flares: report of the OMERACT 11 RA flare Workshop. J Rheumatol. 2014;41(4):799-809.

14. Berthelot JM, De Bandt M, Morel J, Benatig F, Constantin A, Gaudin P, et al. A tool to identify recent or present rheumatoid arthritis flare from both patient and physician perspectives: the 'FLARE' instrument. Ann Rheum Dis 2012;71(7):1110-6.

15. Berthelot JM, Preiss P, Langiller M, Guillemin F, Fautrel B. Frequency, severity, and duration of transient flares in rheumatoid arthritis: A survey of 403 patients. Ann Rheum Dis 2014;73.

16. Myasoedova E, Crowson CS, Davis JM, Gabriel SE, Matteson EL. Identifying flare in rheumatoid arthritis: What is the threshold? Arthritis Rheum 2016;68:3251-3.

17. Brunner HI, Lovell DJ, Finck BK, Giannini EH. Preliminary definition of disease flare in juvenile rheumatoid arthritis. J Rheumatol. 2002;29(5):1058-64.

18. Schumacher HR, Edwards LN, Perez-Ruiz F, Becker M, Chen LX, Furst DE, et al. Outcome measures for acute and chronic gout. J Rheumatol. 2005;32(12):2452-5.

19. Schumacher HR, Taylor W, Joseph-Ridge N, Perez-Ruiz F, Chen LX, Schlesinger N, et al. Outcome evaluations in gout. J Rheumatol. 2007;34(6):1381-5.

20. Schumacher HR, Taylor W, Edwards L, Grainger R, Schlesinger N, Dalbeth N, et al. Outcome domains for studies of acute and chronic gout. J Rheumatol. 2009;36(10):2342-5.

21. Taylor WJ, Shewchuk R, Saag KG, Schumacher HR, Jr., Singh JA, Grainger R, et al. Toward a valid definition of gout flare: results of consensus exercises using Delphi methodology and cognitive mapping. Arthritis Rheum 2009;61(4):535-43.

22. Gaffo AL, Schumacher HR, Saag KG, Taylor WJ, Dinnella J, Outman R, et al. Developing a provisional definition of flare in patients with established gout. Arthritis Rheum 2012;64(5):1508-17.

23. MacFarlane LA, Liu C-C, Solomon DH, Kim SC. Validation of claims-based algorithms for gout flares. Pharmacoepidemiol Drug Saf 2016;25(7):820-6.

24. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. Lupus. 1999;8(8):685-91.

25. Petri M, Genovese M, Engle E, Hochberg M. Definition, incidence, and clinical description of flare in systemic lupus erythematosus: A prospective cohort study. Arthritis Rheum. 1991;34(8):937-44.

26. Gladman DD, Urowitz MB, Kagal A, Hallett D. Accurately describing changes in disease activity in systemic lupus erythematosus. J Rheumatol. 2000;27(2):377-9.

27. Ehrenstein MR, Conroy SE, Heath J, Latchman DS, Isenberg DA. The occurrence, nature and distribution of flares in a cohort of patients with systemic lupus erythematosus: A rheumatological view. Rheumatol. 1995;34(3):257-60.

28. Gordon C, Sutcliffe N, Skan J, Stoll T, Isenberg DA. Definition and treatment of lupus flares measured by the BILAG index. Rheumatology. 2003;42(11):1372-9.

29. Petri M, Buyon J, Skoyron M, Kim M. Reliability of SELENA SLEDAI and flare as clinical trial outcome measures [abstract]. Arthritis Rheum. 1997;40(No 9 Suppl: S218.).

30. Ruperto N, Hanrahan LM, Alarcon GS, Belmont HM, Brey RL, Brunetta P, et al. International consensus for a definition of disease flare in lupus. Lupus. 2011;20(5):453-62.

31. Brunner HI, Klein-Gitelman MS, Higgins GC, Lapidus SK, Levy DM, Eberhard A, et al. Toward the development of criteria for global flares in juvenile systemic lupus erythematosus. Arthritis Care Res. 2010;62(6):811-20.

32. Moverley AR, Vinall-Collier KA, Helliwell PS. It's not just the joints, it's the whole thing: qualitative analysis of patients' experience of flare in psoriatic arthritis. Rheumatology. 2015;54(8):1448-53.

33. Moverley AR, Waxman R, De Wit M, Parkinson A, Campbell W, Brooke M, et al. Development of a flare instrument for use in psoriatic disease: A Report from the 2015 GRAPPA Annual Meeting. J Rheumatol. 2016;43(5):974-8.

34. Marty M, Hilliquin P, Rozenberg S, Valat JP, Vignon E, Coste P, et al. Validation of the KOFUS (Knee Osteoarthritis Flare-Ups Score). Joint Bone Spine. 2009;76(3):268-72.

35. Brophy S, Calin A. Definition of disease flare in ankylosing spondylitis: the patients' perspective. J Rheumatol. 2002;29(5):954-8.

36. Gossec L, Portier A, Landewé R, Etcheto A, Navarro-Compán V, Kroon F, et al. Preliminary definitions of 'flare' in axial spondyloarthritis, based on pain, BASDAI and ASDAS-CRP: an ASAS initiative. Ann Rheum Dis. 2016;75(6):991-6.

37. Godfrin-Valnet M, Puyraveau M, Prati C, Wendling D. Flare in spondyloarthritis: Thresholds of disease activity variations. Joint Bone Spine. 2015;82(3):192-5.

38. Vincent A, Whipple MO, Rhudy LM. Fibromyalgia flares: A qualitative analysis. Pain Med. 2016;17(3):463-8.

39. Young AE, Wasiak R, Phillips L, Gross DP. Workers' perspectives on low back pain recurrence: "it comes and goes and comes and goes, but it's always there". Pain. 2011;152(1):204-11.

40. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995;38(1):44-8.

41. Bykerk VP, Bingham ICO, Choy EH, Boire G, Haraoui BP, Lin D, et al. The omeract preliminary flare questionnaire (PFQ) is responsive to change and able to detect clinically important worsening indicating need for treatment change in the canadian early arthritis cohort. Ann Rheum Dis. 2014;73.

42. Van Der Maas A, Van Herwaarden N, Woodworth T, Minten M, Furst D, Christensen R, et al. Validity of omeract preliminary flare questions in a randomized controlled trial, that assesses impact of disease activity guided down-titration of anti-TNF treatment in rheumatoid arthritis patients in low disease activity. Ann Rheum Dis. 2014;73.

43. 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;42(11):2185-9.

44. Lie E, Woodworth TG, Christensen R, Kvien TK, Bykerk V, Furst DE, et al. Validation of OMERACT preliminary rheumatoid arthritis flare domains in the NOR-DMARD study. Ann Rheum Dis. 2014;73(10):1781-7.

45. Bartlett S, Bingham C, Choy E, Xiong J, Boire G, Haraoui B, et al. Assessing significant flares in rheumatoid arthritis: Validity of the outcome measures in rheumatology preliminary flare questions in the Canadian early arthritis cohort. J Rheumatol. 2014;41(7):1492.

46. Bykerk VP, Bingham CO, Choy E, Xiong J, Boire G, Hitchon CA, et al. Patient self-assessments and selected patient reported outcomes may reliably identify rheumatoid disease flare in early rheumatoid arthritis patients. Arthritis Rheum. 2012;64:S887-S8.

47. Morel J, Berthelot JM, Constantin A, Debandt M, Gaudin P, Vittecoq O, et al. The flare-RA (flare in rheumatoid arthritis) questionnaire is able to detect disease activity increase, I.E., Flare, occurring between 2 visits to the rheumatologist. Ann Rheum Dis. 2014;73.

48. Myasoedova E, Crowson CS, Fautrel B, Guillemin F, Matteson EL, Gabriel SE. Identifying Fare in Rheumatoid Arthritis (RA): Performance of the Flare-Assessment in RA (FLARE) Questionnaire in a US Population. Arthritis Rheum 2014;66:S465.

49. de Thurah A, Maribo T, Stengaard-Pedersen K. Patient self-assessment of flare in rheumatoid arthritis: criterion and concurrent validity of the Flare instrument. Clin Rheumatol. 2016;35(2):467-71.

50. Fautrel B, Morel J, Berthelot JM, Constantin A, De Bandt M, Gaudin P, et al. Validation of FLARE-RA, a Self-Administered Tool to Detect Recent or Current Rheumatoid Arthritis Flare. Arthritis Rheumatol. 2017;69(2):309-19.

51. Lizarraga A, Landi M, Schneeberger E, Gallino Yanzi J, Betancur G, Zaffarana C, et al. Validation of the FLARE Questionnaire for the Detection of a Disease Flare in Rheumatoid Arthritis [abstract]. Arthritis Rheumatol. 2016;68(suppl 10).

52. Maribo T, de Thurah A, Stengaard-Pedersen K. Patient-self assessment of flare in rheumatoid arthritis: translation and reliability of the Flare instrument. Clin Rheumatol. 2016;35(4):1053-8.

53. Dougados M, Huizinga TWJ, Choy EH, Bingham CO, 3rd, Aassi M, Bernasconi C. Evaluation of the Disease Activity Score in Twenty-Eight Joints-Based Flare Definitions in Rheumatoid Arthritis: Data From a Three-Year Clinical Trial. Arthritis Care Res. 2015;67(12):1762-6.

54. Portier A, Gossec L, Tubach F, Alfaiate T, Pham T, Saraux A, et al. Patient-perceived flares in rheumatoid arthritis: A sub-analysis of the STRASS treatment tapering strategy trial. Joint Bone Spine. 2016.

55. Teoh N, Gamble G, Horne A, Taylor WJ, Palmano K, Dalbeth N. Mapping the topography of gout flares: Solutions for flare reporting in gout clinical trials. Arthritis Rheum. 2016;68:2996-7.

56. FitzGerald JD, Grossman JM. Validity and reliability of retrospective assessment of disease activity and flare in observational cohorts of lupus patients. Lupus. 1999;8(8):638-44.

57. Isenberg DA, Allen E, Farewell V, D'Cruz D, Alarcón GS, Aranow C, et al. An assessment of disease flare in patients with systemic lupus erythematosus: A comparison of BILAG 2004 and the flare version of SELENA. Ann Rheum Dis. 2011;70(1):54-9.

58. Petri M, Merrill JT, Maciuca R, Davis Jr JC, Kennedy WP. Validation of the revised SELENA flare index in systemic lupus erythematosus. Arthritis Rheum 2013;65:S1085.

59. Thanou A, Chakravarty E, James JA, Merrill JT. How should lupus flares be measured? Deconstruction of the safety of estrogen in lupus erythematosus national assessment-systemic lupus erythematosus disease activity index flare index. Rheumatology. 2014;53(12):2175-81.

60. Murphy SL, Lyden AK, Kratz AL, Fritz H, Williams DA, Clauw DJ, et al. Characterizing pain flares from the perspective of Individuals with symptomatic knee osteoarthritis. Arthritis Care Res 2015;67(8):1103-11.

61. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. BMJ. 2015;350:g7818-g.

62. Stanton TR, Latimer J, Maher CG, Hancock M. Definitions of recurrence of an episode of low back pain: a systematic review. Spine. 2009;34(9):E316-E22.

Table 1: Themes used in definition of flare

Themes	Examples
Worsening of condition (WC)	Worsening of disease activity/condition, a cluster of symptoms, new or worse signs, increase in disease activity.
Change in treatment (CT)	(Re)initiation or change (increase, change medication) of therapy, increase/add self-management.
Pain (P)	Pain intensity, nocturnal pain/awakenings, DAS28.
Symptom intensity (SI)	Intensity of symptoms, overwhelming physical symptoms.
Duration (D)	Duration, persistent, time to maximum pain level, time to complete resolution of pain, duration varying between days and weeks.
Impact on function (IF)	Function, participation, physical symptoms, changes in daily activity, Childhood Health Assessment Questionnaire.
Joint symptoms (JS)	Number of joints with active arthritis/limited range of motion, swollen/warm/tender/stiff/painful joints, DAS28.
Patient self-reported state (PS)	Self-reported flare, patient global assessment, parent and patient global assessment of overall well-being.
Biomarkers (B)	Laboratory tests (e.g. ESR, CRP levels, acute phase marker).
Emotional symptoms (ES)	Emotional/psychological changes/symptoms/consequences), health related quality of life, coping, anger, depression, withdrawn.
Physician assessment (PA)	Physician global assessment.
Fatigue (F)	Physical fatigue, emotional fatigue.
Other (O)	Frequent, rare and random, manageable, flu-like symptoms/fever, muscle spasm, cramp, burning or tightness in the muscle, sweats, loss of appetite, grey pallor, shortness of breath, throughout the entire body.

Table 2: Extracted data from derivation studies (Group A)

Study	Process	Primary purpose of study	Definition or domains	Themes *	Measure of disease activity/symptom severity used for comparison	Comments
Rheumatoid Arthritis (RA)						
RA OMERACT Special Interest Group						
Bingham et al.[5]	Literature review; focus groups; consensus meeting	Describe flare from patient's perspective; identify domains required for definition	A flare occurs with any worsening of disease activity that would, if persistent, in most cases lead to initiation or change of therapy; ***and*** a flare represents a cluster of symptoms of sufficient duration and intensity to require initiation, change, or increase in therapy	WC, D, CT, SI	DAS44/DAS28	Two versions proposed - "change from an earlier state" or "absolute disease activity state". 120 participants (11 patients)
Alten et al.[7]	Report from OMERACT9	Report of process undertaken by OMERACT	Worsening of disease activity that leads to an assessment for a change in therapy - adapted from Bingham et al.[8]	WC, PA, CT		Acknowledge patients include function and participation, but not included in definition
Bingham et al.[6]	Report from OMERACT10	Report domains identified through consensus	Flare represents a cluster of symptoms of sufficient duration and intensity to require (re)initiation, change, or increase in therapy -	WC, D, SI, CT		Change of wording to "return of/(re)initiation".

			adapted from Bingham et al.[8]		
Hewlett et al.[12]	14 Focus groups across 5 countries	Explore patient perspective of RA flares	Themes highlighted by patients: Individual context, uncertainty, symptoms and early warnings, self-management of intensifying symptoms, uncontrollable, uncertainty and seek help when symptoms cannot be contained	WC, CT, SI.	Patients report current patient global VAS does not capture flare

Study	Process	Primary purpose of study	Definition or domains	Themes*	Measure of disease activity/symptom severity used for comparison	Comments
Bartlett et al.[8]	Delphi	3 phase Delphi to determine RA flare domains	Consensus achieved in 8 domains (pain, function, swollen joints, tender joints, participation, stiffness, patient global assessment and self-management)	P, IF, J, PS, CT		Domains for development of measure of RA flare
Byker et al.[13]	Report of expert and patient consensus meeting	Endorsement of a core domain set to measure RA flare	Pain, function, tender joints, swollen joints, stiffness, patient global assessment, participation and self-management	P, IF, JS, PS, CT		Laboratory values/EGA endorsed by HCP (75%), not patients. Fatigue - endorsed by patients (76%), not HCP. Domains identified by 50-70% of patients - sleep, systemic features and emotional distress
Strategy of treatment in Patients with Rheumatoid Arthritis group of the French Society of Rheumatology						
Berthelot et al.[14]	Delphi	Build a flare questionnaire based on identified domains	Joint swelling, stiffness or pain, night pain, worsening of condition, fatigue, emotional consequences, analgesic intake, changes in daily activity	JS, P,WC, F, ES, CT, IF		Approach encompassing any disease exacerbation (spontaneously regressive or long lasting), more suitable for clinical research and daily practice
Berthelot et al.[15]	Qualitative survey	Assess features of RA transient flares, plus physicians'/patients' attitudes to flares	Joint swelling, stiffness or pain, night pain, worsening of condition, fatigue, emotional consequences, analgesic intake, changes in daily activity	JS, P, WC, F, ES, CT, IF		90% of patients considered transient flares could induce lasting damage, 15% anticipated rheumatology visit

Study	Process	Primary purpose of study	Definition or domains	Themes*	Measure of disease activity/symptom severity used for comparison	Comments
			Independent research group			
Myasoedova et al.[16]	Prospective Study	Establish threshold for RA flare using the FLARE questionnaire in RA	Flare defined as a score of 2.5 or more on The Flare Instrument	JS, P, WC, F, ES, CT, IF	The Flare instrument	
			Juvenile Rheumatoid Arthritis (jRA)			
Brunner et al.[17]	RCT	Create and assess candidate flare criteria	(1) Worsening in any 2/6 CRV by ≥40% without improvement in more than 1 of the remaining CRV by ≥30%; (2) Worsening in 3 of the 6 CRV by ≥ 30%; and (3) any worsening of the Childhood Health Assessment Questionnaire, worsening of ESR by ≥ 30% and worsening of the active joint count by ≥ 10%	WC, JS, PA, PS, B, IF	Core response variables (CRV)	Placebo treatment was criterion standard for disease flare – patients on placebo expected to experience flare vs. no flare with treatment.
			Gout			
Taylor et al.[21]	Delphi process/ Cognitive mapping	Identify domains for standard definition of gout flare	Swollen joints, tender joints, warm joints, patient self-reported of pain, patient self-report global assessment, time to maximum pain level, time to complete resolution of pain, functional status and acute-phase marker	JS, P, PS, D, IF, B		

Study	Process	Primary purpose of study	Definition or domains	Themes*	Measure of disease activity/symptom severity used for comparison	Comments
Gaffo et al.[22]	Cohort Study	Develop definitions for gout flare from patient-reported features	(i) Patient-reported warm joint, (ii) Patient-reported swollen joint,(iii) Patient-reported pain at rest score of >3 (0–10 scale), and (iv) Patient-reported flare	P, JS, PS	Rheumatologist opinion as a gold standard	Best discrimination – 3/more of 4 criteria (sensitivity 91%, specificity 82%). All 4 criteria provided highest specificity (96%), PPV (85%)
MacFarlane et al.[23]	Cohort study	Develop/validate a claims-based algorithm to identify gout flares	ICD-9 plus: medication claim for any gout-related medications/Colchicine/ NSAID/Cox-2 selective inhibitor/oral Glucocorticoids/; ICD-9; CPT or J code within 7 days/CPT or J code on same day	CT, JS, PA		
Systemic Lupus Erythematosus (SLE)						
Petri et al.[25]	Cohort study	Investigate and quantify flares	Change of 0.3 on the LAI and a change ≥3 or on the SLEDAI	JS, P, PS, B, F, O		LAI and SLEDAI include domains related to non-musculoskeletal symptoms

Study	Process	Primary purpose of study	Definition or domains	Themes*	Measure of disease activity/symptom severity used for comparison	Comments
Petri et al.[29]	RCT	Determine reliability of SELENA SLEDAI to identify flares	*Mild/moderate flare*: a change in SLEDAI ≥3 points or new/worse skin, stomatitis, serositis, arthritis, fever or increased prednisone <0.5mg/kg/d or added NSAID/Plaquenil or ≥1 increase in a PGA. *Severe flare*: change in SLEDAI>12 or new/worse CNS-SLE, vasculitis, nephritis, myositis, Plt<60000 hemolyticanemia with Hb<7mg/dl, requiring doubling or > 0.5mg/kg/d prednisone or hospitalization for SLE or prednisone >0.5mg/kg/d or new immunosuppressive or increased physician's global assessment to >2.5	JS, P, PA, CT, B, O		SLEDAI includes domains related to non-musculoskeletal symptoms
Gladman et al.[26]	Cohort Study	Determine which SLEDAI score value correlate with clinician diagnosis of flare	Increase in SLEDAI of >3	JS, PA, P, B, O	SLEDAI	SLEDAI includes domains related to non-musculoskeletal symptoms.

Study	Process	Primary purpose of study	Definition or domains	Themes*	Measure of disease activity/symptom severity used for comparison	Comments
Gordon et al.[28]	Cohort study	Evaluate BILAG score	Moderate disease flare can be defined as a new B score following a C, D, or E score according to the BILAG index	JS, P, B, F, O	BILAG	
Ehrenstein et al.[27]	Cohort Study	Examine occurrence, nature and distribution of flares according to organ involvement	Change in the BILAG index from a score D or E to B, or an increase to an A from any previous score, compared to a previous visit in any of eight organs or systems	JS, P, B, F, O	BILAG	
Ruperto et al.[30]	Consensus Panel	Define SLE flare	Flare is a measurable increase in disease activity in one or more organ systems involving new or worse clinical signs and symptoms and/or laboratory measurements. It must be considered clinically significant by the assessor and usually there would be at least consideration of a change or an increase in treatment	WC, B, PA, CT		Unclear number of patients in Consensus Panel
Juvenile Systemic Lupus Erythematosus (jSLE)						
Brunner et al.[31]	2 international Delphi studies	Define jSLE flares	Flare is a measurable worsening of jSLE disease activity in at least one organ system, involving new or worse signs of disease that may be accompanied by new or worse SLE symptoms. Depending on the severity of the flare, more intensive therapy may be required	WC, SI, CT		North America (n= 72), Europe (n= 35), Central and South America (n=38), Asia (n= 5), Africa (n= 3), Australia (n= 3). 96% consensus reached

Study	Process	Primary purpose of study	Definition or domains	Themes*	Measure of disease activity/symptom severity used for comparison	Comments
			Psoriatic Arthritis (PsA)			
Moverley et al.[32]	Qualitative study	Identify patients' perspectives of PsA flare	An overwhelming collection of physical, psychological and emotional symptoms	SI, IF, ES		Flare duration considered to vary from hours-days
Moverley et al.[33]	2 Delphi surveys; Face-to-face discussion	Develop questionnaire to determine presence of flare in PsA	A change in disease state that necessitates a change in treatment ***or*** as a marked worsening of ability to continue with activities of daily living." Domains: Articular, skin, emotional, participation and fatigue	CT, WC, IF, JS, ES, F		"Flare" - short-lived and acute; "worsening of disease" - slower and longer-lived
			Osteoarthritis (OA)			
Marty et al.[34]	Observational cross-sectional	Derive and validate a definition of flare in OA	A cut-off of 7 on the KOFUS scale	JS, D, P, IF	KOFUS Scale	

Study	Process	Primary purpose of study	Definition or domains	Themes*	Measure of disease activity/symptom severity used for comparison	Comments
			Ankylosing Spondylitis (AS)			
Brophy and Calin[35]	Group Meetings	Examine patient's perception of factors important in defining flare	(1) *Localized to one area*: frequent, duration varying between days and weeks, acute and sudden pain in one area, which may move to other area. Occasional swelling, inability to move joint, immobility, fatigue, bad temper, withdrawn. Attitude: adapt and learn to live with it. (2) *Throughout the entire body*: Rare and random, duration varying between days and weeks, paralysing and throbbing pain inside and in every joint. Immobility, fatigue, muscle spasm, cramp, burning or tightness in the muscle. Sweats, fever, flu-like illness, loss of appetite, grey pallor, shortness of breath,	D, P, JS, IF, F, ES, O		

			depression, anger. Attitude: devastating every time.		
			Assessment of Spondyloarthritis (ASAS) group		
Gossec et al.[36]	Systematic literature review, Vignette exercises	Develop definition for "flare" based on validated composite indices	Candidate flare definitions: Δ pain ≥2 + final value ≥4; Δ pain ≥3; If observed value is ≥4: Δ pain ≥2 points, otherwise: Δ pain ≥3 points; Δ BASDAI ≥2 points; Δ BASDAI ≥2 points + final value ≥4; Δ BASDAI ≥3 points; Δ BASDAI ≥3 points + final value ≥4; If observed value is ≥4, Δ BASDAI ≥2 points, otherwise: Δ BASDAI ≥3 points; Δ ASDAS ≥0.6; Δ ASDAS ≥0.9; Δ ASDAS ≥1.1; Δ ASDAS ≥0.6 + observed ASDAS ≥1.3	P, JS, PS, F, B	ASDAS, BASDAI

Study	Process	Primary purpose of study	Definition or domains	Themes *	Measure of disease activity/symptom severity used for comparison	Comments
Godfrin-Valnet et al.[37]	Longitudinal study	Evaluate thresholds of disease activity variations using validated instruments	Variation of ≥2.1 in BASDAI, 0.8 units in ASDAS-ESR and 1.3 in ASDAS-CRP	P, JS, PS, F, B		
Fibromyalgia						
Vincent et al.[38]	Qualitative Study	Describe flares in patients with fibromyalgia	Pain, flu-like symptoms, fatigue and emotional symptoms	P, ES, F, O		Patients argued flares differed from everyday symptoms by being larger and more intense than usual, to the point where it feels disabling
Low back pain (LBP)						
Young et al.[39]	Qualitative Study	Investigate patients' perspectives of recurrent LBP episodes	An increased experience of pain, but not too intense – causing some additional activity limitations and participation restrictions, but manageable	P, SI, IF, O		

DAS44/DAS28 – Disease Activity Score. The numbers 44 and 28 refers to the 44/28 joints that are examined in this instrument; VAS - Visual Analogue Scale; EGA - Evaluator Global Assessment; HCP – Health Care Professionals; RCT – Randomized controlled trial; CRV – Core Response Variables; PGA – Physician's Global Assessment; ESR – Erythrocyte Sedimentation Rate; ICD-9 – International Classification of Disease-9; NSAID - Nonsteroidal Anti-Inflammatory Drugs; CPT – Current Procedural Terminology; J code - Healthcare Common Procedure Coding System (HCPCS) codes for injection of drugs; LAI – Lupus Activity Index; SLEDAI - Systemic Lupus Erythematosus Disease Activity Index; SELENA Flare tool - Safety of Estrogens in Lupus Erythematosus National Assessment Flare tool; CNS-SLE – Central Nervous System-Systemic Lupus Erythematosus; PLT - Platelet; HB - Haemoglobin; BILAG - British Isles Lupus Assessment Group 2004 index; KOFUS – Knee Osteoarthritis Flare-ups Score; ASDAS - Ankylosing Spondylitis Disease Activity Score;

ASDAS-CRP – Ankylosing Spondylitis Disease Activity Score-C-reactive protein; ASDAS-ESR - Ankylosing Spondylitis Disease Activity Score-Erythrocyte Sedimentation Rate; BASDAI – Bath Ankylosing Spondylitis Disease Activity Index.
* See Table 1 for abbreviations for Themes

Table 3: Extracted data from validation studies (Group B)

Study	Sample characteristics	Validation method d	Definition/domain investigated	Comparator for Validation	Outcomes of validation
		Rheumatoid Arthritis (RA)			
		RA OMERACT Special Interest Group			
Bykerk et al.[41]	N=501, observational cohort study	Longitudinal Observational Study	PFQ	Rheumatologist diagnosis of flare, patient self-reported flare, and 2 proposed DAS28-based criteria	PFQ domain scores changed with flare status of patients and remained stable in patients with unchanged flare status. Inter-rater reliability - flare status of rheumatologist vs. patient - 74% (Kappa 0.34); lower with most specific DAS criterion - 72% (Kappa 0.17)
Van Der Maas et al.[42]	N=180, treated with usual care or TNF-blockers for 6 months, DAS28<3.2, RA treatment >6 months	Randomized Controlled Trial	PFQ	DAS28	Scores for each PFQ domain unchanged in patients without flare, change across all PFQ items in those with flare defined by DAS28
Bartlett et al.[43]	N/A	Report of OMERACT 2014 RA Flare Group and breakout groups	N/A	N/A	Breakout groups - morning stiffness duration may not capture patients' experiences; insufficient scope of self-management activities

Study	Sample characteristics	Validation method d	Definition/domain investigated	Comparator for Validation	Outcomes of validation
Lie et al.[44]	N=1195, taking DMARDS	Longitudinal Observational Study: Validity of domains assessed using logistic regression analysis	**Validation 1:** "Domains of flare" identified by 28-SJC; 28 TJC; ESR, CRP, VAS for pain, fatigue and PGA[1], SF-36, MHAQ score, morning stiffness intensity and duration (BASDAI). **Validation 2:** Patient self-reported worsening, treatment change	DAS28 + three working definitions: patient self-reported worsening, treatment change, or both self-reported worsening and treatment change DAS28	"Domains of flare" more strongly correlated with "self-reported worsening of symptoms" than "change in treatment". "Domains of" flare were related to DAS28 DAS28 more related to "patient self-reported worsening" than "change in treatment".
Bartlett et al.[45]	N=1190, from Canadian early Arthritis CoHort (CATCH)	Cohort Study: Discriminant validity of PFQs, convergent validity among PFQs and validated RA measures	PFQs	HAQ, SF12, RADAI, WPAI and Patient Global	Agreement among single item PFQs and validated RA measures. PFQ scores higher across domains for patients reporting flares
Bykerk et al.[46]	N=512, CATCH	Cohort Study: Agreement between patients and rheumatologists	Pain, function, stiffness, participation, coping, patient global assessment, fatigue and self-management according to *patients*	Pain, function, stiffness, participation, coping, patient global assessment, fatigue and self-management according to *rheumatologists*	Modest agreement - patients vs. rheumatologists. Pain, function, stiffness, coping, participation, fatigue discriminate patients with vs. without flare

Study	Sample characteristics	Validation method d	Definition/domain investigated	Comparator for Validation	Outcomes of validation
Strategy of treatment in Patients with Rheumatoid Arthritis group of the French Society of Rheumatology					
Fautrel et al.[50]	N=138, ≥18 years, (1987 ACR and/or 2010 ACR/EULAR	Prospective Study	FLARE-RA self-administered questionnaire*	RAPID-3, RAID, DAS28 and HAQ	Good content and construct validity of FLARE-RA. Floor effect, no ceiling effect. FLARE-RA total score correlated with DAS28 (r=50.63, P<0.001), RAID (r=50.80, P<0.001), RAPID-3 (r=50.77, P<0.001), and HAQ (r=50.53, P<0.001). ICC(reliability) - 0.94 (95% CI 0.92–0.96)
Morel et al.[47]	N=132, 13 centres (1987 ACR criteria, duration >6 months), treated DMARDs	Prospective Study	FLARE-RA questionnaire*	RAPID3 self-administered questionnaire	FLARE-RA questionnaire detected past flares defined by RAPID3
Myasoedova et al.[48]	N=160 (1987 ACR criteria)	Cohort Study: Correlation - FLARE questionnaire vs. clinical/laboratory measures	FLARE questionnaire*	BRAF, HAQ, VAS pain, CRP, IL6, PCA, PGA[1]	FLARE score/subscales correlated with all comparators, except PCA CRP correlated with overall score and systemic subscale

Study	Sample characteristics	Validation method d	Definition/domain investigated	Comparator for Validation	Outcomes of validation
de Thurah et al.[49]	N=117, (ACR 1987/2010 criteria), visiting outpatient clinic, treated with DMARDs	Cohort Study: ROC curves, sensitivity, specificity, PPV and NPV, likelihood ratio and Spearman correlation coefficient	Danish version of Flare instrument (FI)	DAS28-CPR (criterion standard), HAQ and CRP (concurrent validity)	Sensitivity (detect flare identified by DAS28) - 85.4% (95% CI, 72.2; 93.9); specificity - 50.7% (95% CI, 38.4; 63.0). PPV - 53.6% (95% CI, 47.0; 60.1) NPV - 83.9% (95% CI, 7; 91.5). AUC - ~77%. Moderate correlation DAS28-CPR vs. FI, poor correlation CRP vs. FI and FI vs. HAQ
Maribo et al.[52]	N=117, Dutch patients (ACR 1987/2010 criteria) visiting outpatient clinic, treated with DMARDs	Cross-cultural Study: Forward and back-ward translation and calculation of LOA, SEM, MDC and ICC.	Flare instrument*	DAS28	Instrument identifies >80% RA patients without flare
Lizarraga A et al. [51]	N=105, ≥18 years (ACR/EULAR 2010 criteria)	Translation and cultural adaptation	FLARE Questionnaire	DAS28, RAPID, patient's opinion and physician's opinion	FLARE questionnaire validated for detecting RA disease flare
Dougados et al.[53]	N=553 from a RCT	Post hoc analysis of RCT	Increase in DAS28, between 2 visits, of >1.2 or >0.6 if the current DAS28 ≥3.2	Worsening of disease activity which required increased treatment beyond permitted supportive therapy	DAS28-2 - sensitive (88–100%), poor specificity (57–65%), for detecting flare

Study	Sample characteristics	Validation method d	Definition/domain investigated	Comparator for Validation	Outcomes of validation
Portier et al.[54]	N=134, ≥18 years (ACR 1987 criteria)	Post-hoc analysis of Randomized Controlled Trial	Increase of DAS28>0.6 and an absolute value of DAS28>2.6 (physician assessment)	Patient-reported flares	Flare concept differs between patients and physicians (Kappa 0.44). Patients consider flare is more than disease activity
			Gout		
Teoh et al.[55]	N=120 from RCT	Randomized Controlled Trial: Correlation between flare reporting and measures of gout activity	Pain score >3, patient self-report flare (SRF), flare count, time to first flare, number of days with (SRF) and (SRF) requiring medication	Area under curve (AUC) variable time plot analysis, patient and physician global assessments, joint counts, and CRP	Flare reporting flare correlated with measures of disease severity, except time to first flare
			Systemic Lupus Erythematosus (SLE)		
FitzGerald and Grossman.[5] [6]	N=22 - intra-rater reliability; N=26 inter-rater reliability	Cohort Study: Validity, intra- and inter-rater reliability between two physicians abstracting PGA^2, SLEDAI and SELENA flare tool	Flare defined by PGA^2, SLEDAI and SELENA Flare tool	Flare defined by PGA^2, SLEDAI and SELENA Flare tool	Poor agreement flares definitions among instruments. Poor validity of timing of flare at a specific patient-level. Flare better assessed over specific time period
Gladman et al.[26]	N=230	Cohort Study: Descriptive statistics identified SLEDAI score = flare defined by physicians	SLEDAI scores	Physicians' impression of disease activity (new or increased therapy for activity disease, expression of concern, use of "flare" in notes)	Increased SLEDAI >3 = flare identified by physician. Improvement - reduced SLEDAI >3; persistently active disease - increase/decrease of SLEDAI up to 3; remission - SLEDAI = 0

Study	Sample characteristics	Validation method d	Definition/domain investigated	Comparator for Validation	Outcomes of validation
Isenberg et al.[57]	N=16	Cohort Study: Assessment of internal reliability of 3 instruments (ICC with 95% CI)	BILAG 2004, SELENA Flare and PGA^2 (LFA Flare definition used by 16 assessors)	BILAG 2004, SELENA Flare and PGA^2 (LFA Flare definition used by 16 assessors)	BILAG 2004 defined flares performed better [ICC (95% CI) at 0.54 (0.32 to 0.78)] than SELENA flare [at 0.21 (0.08 to 0.48)] and PGA^2 tools [0.18 (0.06 to 0.45)]
Petri et al.[58]	N=235, moderate to severe disease activity using rontalizumab (anti-interferon alpha)	Randomized Controlled Trial: rSFI vs. SFI and BILAG flare index. Inter instrument agreement - Cohen's kappa/weighted kappa coefficient	rSFI	SFI, BILAG flare index	Substantial agreement between rSFI and SFI. BILAG flare index - better agreement with rSFI than SFI
Thanou et al.[59]	N=91 (1997 modified ACR criteria for SLE)	Cohort Study: cSFI and eSFI were compared. Descriptive statistics used to describe measures of disease activity and flare	cSFI (includes medication)	eSFI (excludes medication), SLEDAI, BILAG, PGA^2	eSFI better than cSFI eSFI- improves discrimination of mild vs. moderate flares, selects more ill patients with clinical worsening for each category of flare Medication - not distinguish the severity of a clinical flare

Study	Sample characteristics	Validation method d	Definition/domain investigated	Comparator for Validation	Outcomes of validation
Juvenile Lupus Erythematosus Systemic (jSLE)					
Brunner et al.[31]	98 Children with jSLE, ≤16 years	Multicentre Study	A flare is a measurable worsening of juvenile SLE disease activity in at least one organ system, involving new or worse signs of disease that may be accompanied by new or worse SLE symptoms. Depending on the severity of the flare, more intensive therapy may be required	Rate change in the jSLE course by paediatric rheumatologist (10+ year experience)	jSLE flares determined by Physician-rated disease activity, disease activity index score, health related quality of life
Osteoarthritis (OA)					
Murphy et al.[60]	Symptomatic knee OA (n=45)	Cohort Study: % patient's agreement with investigator's flare definition	Inadequate pain relief for an episode of intense pain that is usually brought on by too much activity	Patients' definition	~50% participants - definition did not match personal definition. 11% participants "very much" agreed
Marty et al.[34]	641 patients (285 – flare, 356 – stable condition)	CS: Sensitivity, specificity, PPV and NPV	KOFUS (cut-off score=7)	Rheumatologist diagnosis	Sensitivity - 87.0%, specificity - 87.9%, PPV 85.8%, NPV 89.0%

PFQ – Preliminary Flare Questionnaire; DAS/DAS28 – Disease Activity Score. The number 28 refers to the 28 joints that are examined in this instrument; DAS28-CRP - A composite score including the patient's global assessment, report of physical functioning (HAQ), and the measurement of an acute phase reactant (CRP), together with a physician-based count of tender and swollen joints; N/A – Not applicable; DMARDS – Disease Modifying Anti-rheumatic Drugs; 28-SJC – 28-Swollen Joint Counts; 28-TJC – 28-Tender Joint Counts; ESR – Erythrocyte Sedimentation Rate; CRP – C-reactive protein; VAS - Visual Analogue Scale; PGA^1 - Patient Global Assessment; SF36 – 36-item short form survey (SF36 Bodily pain, SF36 Physical functioning, SF36 Social functioning, SF36 Role limitations physical, SF36 Role limitations emotional, SF36 Mental health); MHAQ – Modified Health Assessment Questionnaire; BASDAI – Bath Ankylosing Spondylitis Disease Activity Index; HAQ - Health Assessment Questionnaire; SF12 – Short Form 12 Questionnaire; ACR criteria – American College of Rheumatology criteria; RAID - Rheumatoid Arthritis Impact of Disease; RAPID - Routine Assessment of Patient Index Data; RAPID3 – Routine Assessment of Patient Index Data 3; BRAF - Bristol Rheumatoid Arthritis Fatigue; ICC – Intra-class Correlation Coefficients; IL6 - Interleukin-6; ROC – Receiver Operating Characteristic; PPV – Positive Predictive Value; NPV – Negative Predictive Value; PCA - Physician Clinical Assessment; SEM – Standard Error of the Measurement; Work productivity and Activity Impairment Questionnaire – WPAI; LOA – Limits of Agreement; MDC – Minimally Detectable Change; PGA^2 – Physician's Global Assessment; LFA – Lupus Foundation of America; SLEDAI - Systemic Lupus Erythematosus Disease Activity Index; SELENA Flare tool - Safety of Estrogens in Lupus Erythematosus National Assessment Flare tool; BILAG - British Isles Lupus Assessment Group 2004 index; SFI - Selena Flare Index; rSFI – Revised Selena Flare Index; eSFI – Experimental Selena Flare Index; cSFI – Classic Selena Flare Index; jSLE – Juvenile Systemic Lupus Erythematosus; OA – Osteoarthritis; KOFUS – Knee Osteoarthritis Flare-ups Score.

* Flare self-reported questionnaire, Flare-RA questionnaire, Flare assessment in RA (FLARE) questionnaire and Flare Instrument (FI): all refer to the same instrument.

Table 4: Assessment of quality of derivation process

Study	*Clear identification of condition*	*Perspective: Patient*	*Perspective: Clinician*	*Perspective: Experts*	*Derivation Process**	*Used a research method to identify a threshold/cut-off for flare*	*Consensus Process*
Rheumatoid Arthritis (RA)							
RA OMERACT Special Interest Group							
Bingham et al.[5]	Yes	Yes	Yes	Yes	LR, F, CM, W	No	Yes
Alten et al.[7]	Yes	N/A	N/A	N/A	Q, RCT, L	No	Yes
Bingham et al.[6]	Yes	N/A	N/A	N/A	Q, F, D, LR, RCT, B	No	Yes
Hewlett et al.[12]	Yes	Yes	No	No	F	No	Yes
Bartlett et al.[8]	Yes	Yes	Yes	Yes	D	No	Yes
Bykerk et al.[13]	Yes	Yes	Yes	Yes	CM	No	Yes
Strategy of treatment in Patients with Rheumatoid Arthritis group of the French Society of Rheumatology							
Berthelot et al.[14]	Yes	Yes	Yes	Yes	Q, D	No	Yes
Berthelot et al.[15]	Yes	Yes	Yes	No	Q	No	Yes
Independent research group							
Myasoedova et al.[16]	Yes	Yes	Yes	No	P	Yes	No
Juvenile Rheumatoid Arthritis (jRA)							
Brunner et al.[17]	Yes	No	Yes	No	RCT	Yes	No
Gout							
Taylor et al.[21]	Yes	Yes	Yes	Yes	D	No	Yes
Gaffo et al.[22]	Yes	Yes	Yes	Yes	C	No	No
MacFarlane et al.[23]	Yes	No	Yes	No	C	No	No
Systemic Lupus Erythematosus (SLE)							
Petri et al.[25]	Yes	No	Yes	No	C	No	No
Petri et al.[29]	Yes	No	Yes	No	RCT	Yes	No
Gladman et al.[26]	Yes	No	Yes	No	C	Yes	No
Gordon et al.[28]	Yes	No	Yes	No	C	Yes	No
Ehrenstein et al.[27]	Yes	No	Yes	No	C	No	No
Ruperto et al.[30]	Yes	Yes	Yes	Yes	CP	No	Yes
Study	***Clear***	***Perspecti***	***Perspecti***	***Perspecti***	***Derivat***	***Used a***	***Consen***

	identification of condition	*ve: Patient*	*ve: Clinician*	*ve: Experts*	*ion Process **	*research method to identify a threshold/cut-off for flare*	*sus Process*
Juvenile Lupus Erythematosus Systemic (jSLE)							
Brunner et al.[31]	Yes	No	Yes	No	D	No	Yes
Psoriatic Arthritis (PsA)							
Moverley et al.[32]	Yes	**Yes**	No	No	Q	No	No
Moverley et al.[33]	Yes	Yes	Yes	No	D, B	No	Yes
Osteoarthritis (OA)							
Marty et al.[34]	Yes	Yes	Yes	No	OCS	Yes	No
Ankylosing Spondylitis (AS)							
Brophy and Calin[35]	Yes	Yes	No	No	GM	No	No
Assessment of Spondyloarthritis (ASAS) group							
Gossec et al.[36]	Yes	No	Yes	Yes	LR, V	No	No
Godfrin-Valnet et al.[37]	Yes	Yes	Yes	No	L	Yes	No
Fibromyalgia							
Vincent et al.[38]	Yes	Yes	No	No	Q	No	No
Low back pain (LBP)							
Young et al.[39]	Yes	Yes	No	No	Q	No	No

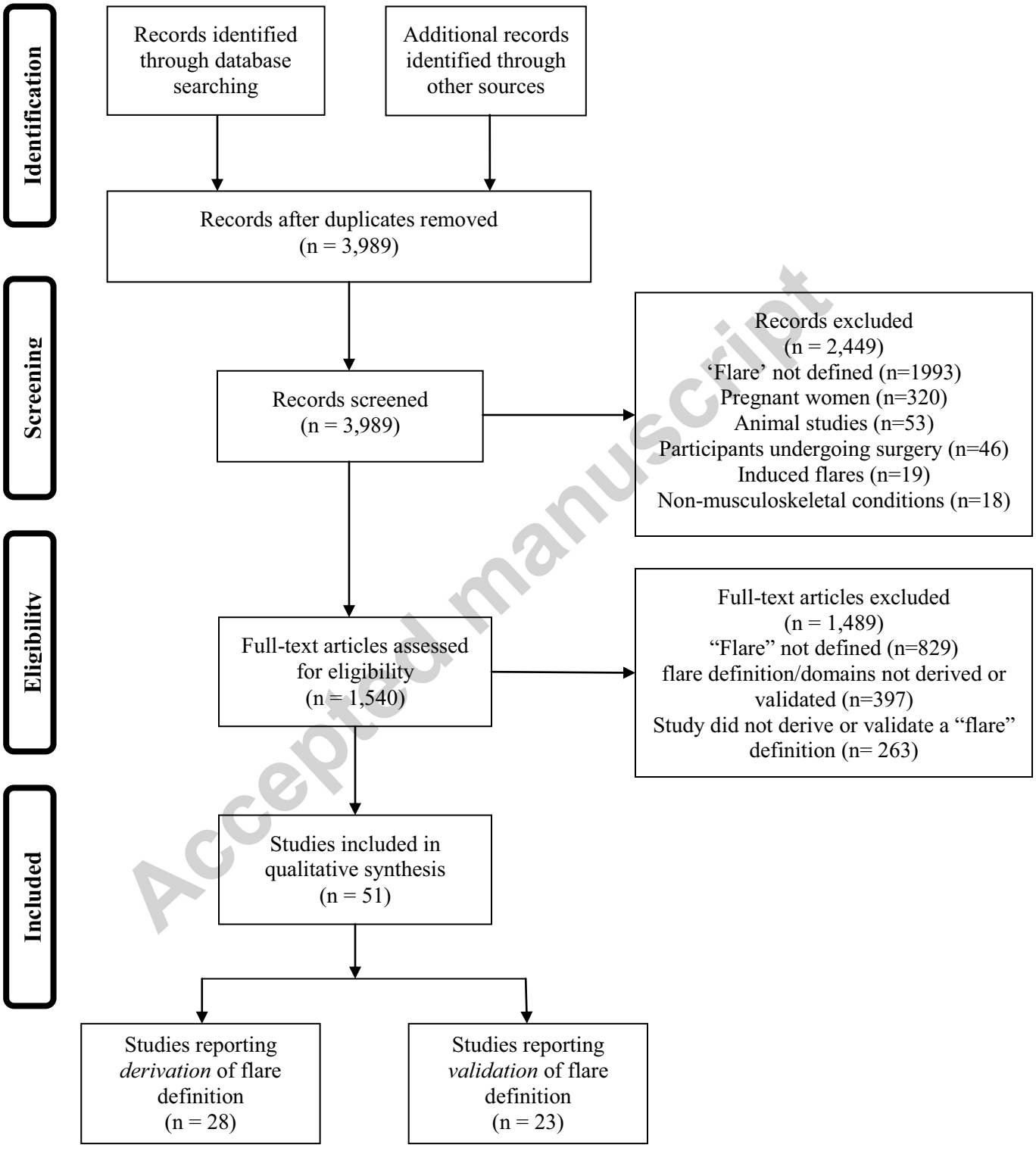
LR – Literature review; F – Focus group; CM – Consensus Meeting; W – Workshop; Q – Qualitative Study; RCT – Randomised Controlled Trial; D – Delphi; B – Breakout groups; P – Prospective Study; C – Cohort study; CP – Consensus Panel; OCS – Observational cross-sectional study; GM – Group meetings; V – Vignette exercise; L – Longitudinal study.
*All definitions derived comprised multiple domains.

Table 5: Assessment of quality of validation process

Study	*Clear identification of condition*	*Representative sample*	*Clear description of comparator*	*Blinding of patients and clinicians to results of the comparator*	*Inclusion of follow-up*
Rheumatoid Arthritis (RA)					
RA OMERACT Special Interest Group					
Bykerk et al.[41]	Yes	Patients from a Canadian early ArThritis CoHort (Observational Study)	Yes	Unclear	Yes
Van Der Maas et al.[42]	Yes	RCT using TNF-blockers or usual care	Yes	Yes	Yes
Bartlett et al.[43]	Yes	N/A	N/A	N/A	N/A
Lie et al.[44]	Yes	5 Norwegian rheumatology departments using DMARDS and/or biologics.	Yes	Yes	Yes
Bartlett et al.[45]	Yes	Patients from a Canadian early ArThritis CoHort (Observational Study)	Yes	Yes	Yes
Bykerk et al.[46]	Yes	Patients from a Canadian early ArThritis CoHort (Observational Study)	Yes	Yes	Yes
Strategy of treatment in Patients with Rheumatoid Arthritis group of the French Society of Rheumatology					
Fautrel et al.[50]	Yes	138 patients >18 years with RA diagnosis (1987 ACR and/or 2010 ACR/EULAR criteria), disease evolving for ≥6 months, treatment with synthetic or biologic DMARDS for ≥2 months and stable symptomatic treatment for ≥2months	Yes	Yes	Yes
Morel et al.[47]	Yes	Patients from 13 treated with stable doses of DMARDs	Yes	Yes	Yes
Myasoedova et al.[48]	Yes	160 RA Patients - population-based cohort (1987 ACR criteria)	Yes	Yes	Yes
de Thurah et al.[49]	Yes	Outpatient clinic treated with DMARDs	Yes	N/A	N/A
Maribo et al.[52]	Yes	Out-patient clinic in Denmark	Yes	N/A	N/A
Lizarraga A et al.[51]	Yes	≥18 years, RA (ACR/EULAR 2010 criteria)	Yes	Yes	Yes

Study	*Clear identification of condition*	*Representative sample*	*Clear description of comparator*	*Blinding of patients and clinicians to results of the comparator*	*Inclusion of follow-up*
		Independent Research Groups			
Dougados et al.[53]	Yes	553 patients, 118 centres in 19 countries (Europe, America and Asia) using tocilizumab + methotrexate or placebo	Yes	Yes	Yes
Portier et al.[54]	Yes	≥18 years, RA (ACR 1987 criteria)	Yes	Yes	Yes
		Gout			
Teoh et al.[55]	Yes	120 gout patients	Yes	Yes	Yes
		Systemic Lupus Erythematosus (SLE)			
FitzGerald and Grossman[56]	Yes	Randomly selected from rheumatology outpatient clinics (UCLA – Los Angeles)	Yes	Yes	Yes
Gladman et al.[26]	Yes	Patients from Uni. Toronto Lupus Clinic	No	Yes	Yes
Isenberg et al.[57]	Yes	Limited to patients from 2 lupus clinics seeking care during the previous 3 weeks	Yes	Yes	No
Petri et al.[58]	Yes	Limited to patients from a RCT conducted in USA, Latin America and Europe using rontalizumab	Yes	Yes	Yes
Thanou et al.[59]	Yes	Limited to SLE patients - Oklahoma Lupus Cohort Study, (1997 modified ACR criteria)	Yes	No	Yes
Brunner et al.[31]	Yes	Limited to patients ≤16 years from 7 US paediatric rheumatology clinics	No	Yes	Yes
		Osteoarthritis (OA)			
Murphy et al.[60]	Yes	Limited to moderate to severe pain from the Uni. Michigan Clinics	No	No	Yes
Marty et al.[34]	Yes	Clinical and radiological criteria for knee osteoarthritis (ACR)	No	Yes	No

RCT – Randomized Controlled Trial; DMARDS – Disease Modifying Anti-rheumatic Drugs; ACR– American College of Rheumatology; EULAR - European League Against Rheumatism.



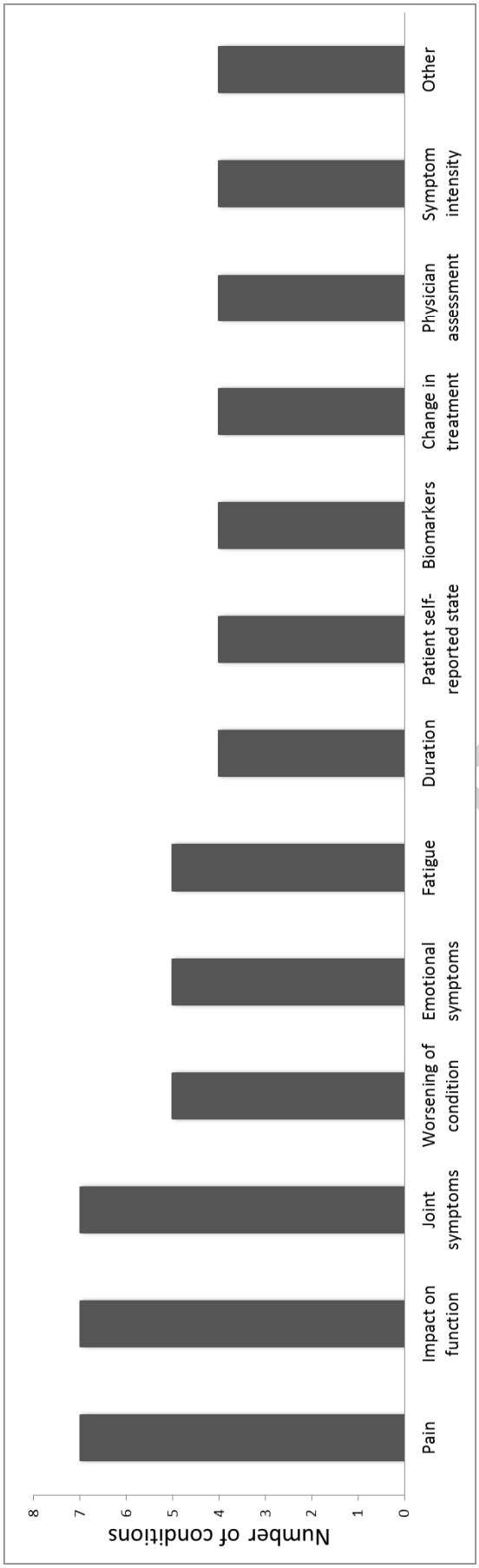


Figure 2. Number of conditions (out of 9 major conditions) that included consideration of each theme in the flare definition.

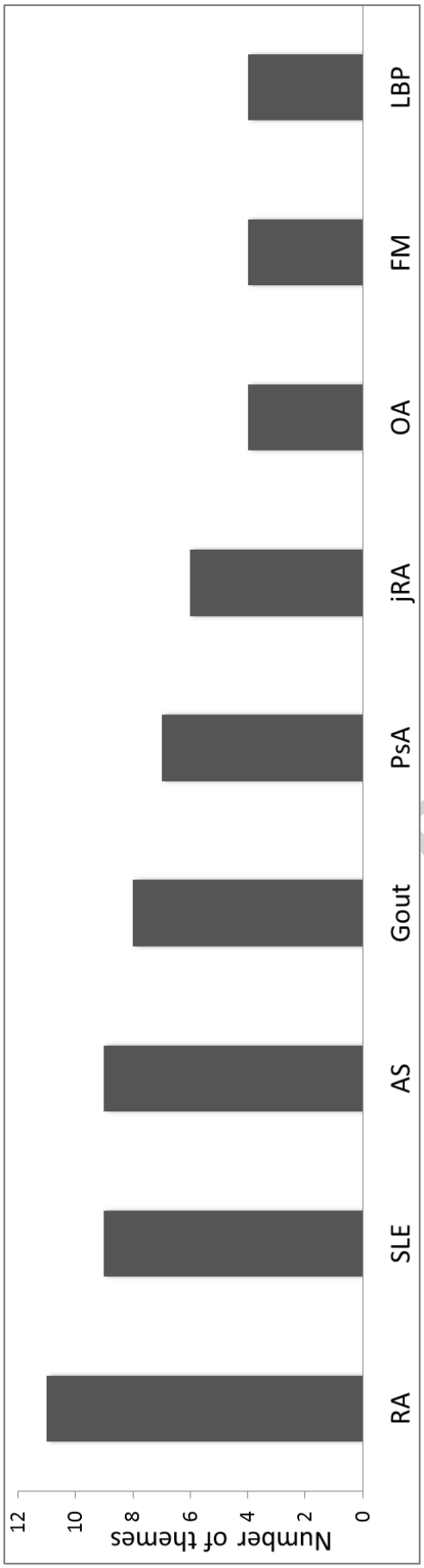


Figure 3. Number of themes considered for each condition (out of 13 themes).