

Identifying Preliminary Domains to Detect and Measure Rheumatoid Arthritis Flares: Report of the OMERACT 10 RA Flare Workshop

CLIFTON O. BINGHAM III, RIEKE ALTEN, SUSAN J. BARTLETT, VIVIAN P. BYKERK, PETER M. BROOKS, ERNEST CHOY, ROBIN CHRISTENSEN, DANIEL E. FURST, SARAH E. HEWLETT, AMYE LEONG, JAMES E. MAY, PAM MONTIE, CHRISTOF POHL, TESSA C. SANDERSON, VIBEKE STRAND, and THASIA G. WOODWORTH, for the OMERACT RA Flare Definition Working Group

ABSTRACT. *Background.* While disease flares in rheumatoid arthritis (RA) are a recognized aspect of the disease process, there is limited formative research to describe them.

Methods. The Outcome Measures in Rheumatology Clinical Trials (OMERACT) RA Flare Definition Working Group is conducting an international research project to understand the specific characteristics and impact of episodic disease worsening, or “flare,” so that outcome measures can be developed or modified to reflect this uncommonly measured, but very real and sometimes disabling RA disease feature. Patient research partners provided critical insights into the multidimensional nature of flare. The perspectives of patients and healthcare and research professionals are being integrated to ensure that any outcome measurement to detect flares fulfills the first OMERACT criteria of Truth. Through an iterative data-driven Delphi process, a preliminary list of key domains has been identified to evaluate flare.

Results. At OMERACT 10, consensus was achieved identifying features of flare in addition to the existing core set for RA, including fatigue, stiffness, symptom persistence, systemic features, and participation. Patient self-report of flare was identified as a component of the research agenda needed to establish criterion validity for a flare definition; this can be used in prospective studies to further evaluate the Discrimination and Feasibility components of the OMERACT filter for a flare outcome measure.

Conclusion. Our work to date has provided better understanding of key aspects of the RA disease process as episodic, potentially disabling disease worsening even when a patient is in low disease activity. It also highlights the importance of developing ways to enhance communication between patients and clinicians and improve the ability to achieve “tight control” of disease. (*J Rheumatol* 2011;38:1751–8; doi:10.3899/jrheum.110401)

Key Indexing Terms:

DISEASE ACTIVITY

RHEUMATOID ARTHRITIS

OUTCOME ASSESSMENT

BACKGROUND

Defining “flare” in RA. Patients with rheumatoid arthritis (RA) and their clinicians recognize and describe periods of

worsening disease activity throughout their disease course. The term “flare” in RA appears to be used to describe periods of disease worsening that encompass a range of symp-

From Johns Hopkins University, Baltimore, Maryland, USA; Schlosspark Klinik, Charité University, Berlin, Germany; McGill University, Montreal, Quebec, Canada; Brigham and Women’s Hospital, Boston, Massachusetts, USA; University of Queensland, Brisbane, Australia; King’s College, London, UK; The Parker Institute, Musculoskeletal Statistics Unit, Copenhagen University Hospital, Frederiksberg, Copenhagen; University of California, Geffen School of Medicine, Los Angeles, California, USA; Department of Nursing and Midwifery, University of the West of England, Bristol, UK; Healthy Motivation, Santa Barbara, California, USA; Arthritis Research Centre of Canada, Chilliwack, British Columbia, Canada; Stanford University, Portola Valley, California, USA; Academic Rheumatology Unit, University of the West of England, Bristol, UK; and Leading Edge Clinical Research LLC, Stuart, Florida, USA.

Supported by UCB, Amgen, BMS, Xoma, Roche/Genentech, Centocor, Pfizer, The Arthritis Society, IDENK Ltd., and OMERACT. COB is partly supported through an ACR-REF “Within Our Reach, Finding a Cure for Rheumatoid Arthritis” grant. RC is sponsored by the Oak Foundation. EC acknowledges financial support from the Department of Health via the

NIHR comprehensive Biomedical Research Centre award to Guy’s and St Thomas’ NHS Foundation Trust in partnership with King’s College London. TGW was formerly an employee of Roche Products Limited, Welwyn Garden, UK.

C.O. Bingham, MD, Johns Hopkins University; R. Alten, MD, Schlosspark Klinik, Charité University; S.J. Bartlett, PhD, McGill University; V.P. Bykerk, MD, Brigham and Women’s Hospital; P.M. Brooks, MD, King’s College; E. Choy, MD, King’s College; R. Christensen, MSc, PhD, Parker Institute; D.E. Furst, MD, University of California, Geffen School of Medicine; S.E. Hewlett, PhD, RN, Department of Nursing and Midwifery, University of the West of England; A. Leong, Healthy Motivation; J.E. May; P. Montie, Arthritis Research Centre of Canada; C. Pohl, MD, Schlosspark Klinik, Charité University; T.C. Sanderson, MD, Academic Rheumatology Unit, University of the West of England; V. Strand, MD, Stanford University; T.G. Woodworth, MD, Leading Edge Clinical Research LLC.

*Address correspondence to Dr. Bingham;
E-mail: clifton.bingham@jhmi.edu*

toms and signs that vary in magnitude and duration; it may lead to a range of actions taken on the part of patients and healthcare professionals (HCP). Some flares are short-lived and self-limited, requiring minimal intervention or requiring nonpharmacologic or other self-management strategies. Some RA flares are more severe or disabling, resulting in a decision to make changes in medical therapy, from initiation or increases of nonsteroidal antiinflammatory agents (NSAID) or corticosteroids, to changes or addition of disease-modifying antirheumatic drugs or biologics. Although a variable disease course, characterized by “good” and “bad” days, is clearly recognized, there is a paucity of research that reflects the experience of disease-worsening from the patient’s perspective. Although patients have identified reducing the number of flares as a desirable goal of therapy, there is limited understanding of the implications of episodic worsening, the domains of importance to reflect a flare state, or the thresholds for detecting meaningful change¹.

Although “flare” is a term commonly used in clinical trials and practice in patients with RA, there are no established standardized criteria to measure flare from either the clinician’s or patient’s perspective. Various flare definitions have been incorporated in RA randomized controlled trials (RCT)², but these have not been uniformly applied or validated as clinically relevant to RA outcomes and prevention of disability. Definitions have included an inverse of response (improvement) criteria such as inverse European League Against Rheumatism (EULAR) response criteria, modifications of American College of Rheumatology (ACR) responder criteria, and increases in swollen and tender joints approximating baseline scores^{3,4,5}. However, the research to document performance characteristics for such outcome measures for deterioration or worsening versus improvement has not been reported. There are data from other rheumatic diseases such as systemic lupus erythematosus that indicate patient-detectable, minimal clinically important differences for worsening occur at lower thresholds than the same measures for improvement^{2,6}. Both for RCT and clinical practice, a standardized flare definition is critical to inform trial design or clinical decision-making. In trial design, it can be an outcome measure, while for clinical practice it can signal a need to change management (e.g., add a treatment, or reinstate drug therapy).

With the advent of effective therapies and treatment strategies that provide states of very low disease activity or remission for many patients, there is a need to consider designing studies that can inform decision-making on modifications of therapy to maintain a low activity state, potentially with less intensive treatment regimens, lowered doses, or altered frequencies of intermittently-dosed medications. A barrier exists, however, in developing RCT study designs that propose to investigate treatment optimization approaches or remission induction/maintenance studies, which

require a definition of disease flare, with “time to flare” or “numbers of flares” as an outcome. Moreover, accurately capturing adverse event data for patients who experience fluctuating disease activity between assessment visits or who withdraw from studies due to RA flare is desirable from a drug safety and regulatory perspective. For longitudinal observational studies (LOS), there is a similar need to understand flare versus expected/acceptable disease oscillations in studying the effectiveness of interventions^{7,8}.

The Outcome Measures in Rheumatology Clinical Trials (OMERACT) RA Flare Definition Working Group, composed of patients and HCP from academia and industry, has been conducting research since OMERACT 8 (2006). They were initially a subgroup of the Drug Safety Working Group⁹, but at OMERACT 9, a Special Interest Group reported initial research documenting the need for a standardized definition of flare, and an anchoring definition was agreed among OMERACT 9 participants^{2,9}. In patient groups, there were compelling indications that the patient experience of flare may not be captured adequately by existing core set measurements used in RCT. For example, patients reported clusters of symptoms and a prodromal state that often preceded the development of overt clinically detectable flare signs^{2,10}. Patients also indicated that, when they sensed a flare was beginning, they engaged in a number of self-management strategies to minimize or potentially avert a major flare. There was consensus that a definition of flare was needed for RCT and LOS, and a recognized need for one that would also be useful in clinical care. There was overwhelming consensus that the patient perspective was essential to the work of the group².

A preliminary operational definition for “Flare” in RA was established:

- A flare occurs with any worsening of (or return of) disease activity that would, if persistent, lead to (re)initiation, or/and change of therapy
- A flare represents a cluster of symptoms of sufficient duration and intensity to require (re)initiation, change, or increase in therapy.

METHODS

The OMERACT 9 Special Interest Group was expanded to include additional interested individuals with expertise in analysis of RCT data, LOS, and qualitative research, with a diverse geographical representation and representatives from multiple stakeholder perspectives including additional patient, healthcare provider, and pharmaceutical research partners. The research agenda developed at OMERACT 9 was refined following the meeting, and moved forward with frequent interactions and integration of findings between the interested subgroups. Teleconferences were conducted every 2 weeks, and face-to-face meetings took place at meetings of the ACR in 2008, EULAR in 2009, and ACR in 2009.

The work processes established were: (1) a patient-focused stream that included multinational qualitative research and Delphi exercises^{10,11,12}; (2) data mining of RCT data; (3) data mining from LOS; and (4) a healthcare provider/researcher work stream including ongoing literature review and Delphi exercises (Figure 1).

Garnering consensus on the definition of disease-worsening in clinical trials was clearly needed, given the proposal and initiation of a growing number of pharmaceutical RCT that incorporated a flare definition. Based on consensus obtained at OMERACT 9 and additional work of the group, there was a decision for a short-term focus toward establishing a working, preliminary flare definition for RCT, but a group recognition that the longer-term goal was to develop a fully validated flare definition across different contexts. While this decision was made, it was found that these activities were not dissociable, but interdependent, such that work moving forward in one area would necessarily inform work in other areas. A second report was published that outlined the additional work of the group and more detailed data analysis plans⁷.

The goals of the Workshop at OMERACT 10 were to present data developed from 4 different work streams, including the results of a patient Delphi exercise based on initial experience in focus groups, and the results of 2 sequential HCP Delphi exercises conducted online. These established consensus on a preliminary core set of domains that captured RA flare from patient and HCP perspectives. This consensus required several working sessions, prior to OMERACT 10, where data were collated from various sources, leading to the presentation of key findings to the OMERACT 10 plenary session. In breakout groups, clinical

and patient research partners moderated discussions of specific areas that needed additional small-group input. Representatives from the breakout groups reassembled and organized their inputs, which were presented and voted upon by the whole assembly. This prioritized relevant domains identified by both patients and HCP, which captured the experience and expression of flare.

In order to evaluate the OMERACT filter of Discrimination, we conducted a statistical analysis of RCT and LOS data to evaluate whether the decision to increase dose or change medication can be reflected by using either single variables or composite indices, alone or together in a model. An observation-based approach, involving analysis of existing RA data for which the clinician considered or categorized the patient as either having a flare or not, was taken from each of the available variables in the RA core set. The ability to detect and discriminate a “flare effect” in study outcomes was evaluated using standardized response means [SRM, i.e., the ratio of the group mean difference to the pooled standard deviation (SD) of the mean change scores], a methodology suitable for detecting the best signal-to-noise ratio (i.e., discriminant capacity). The variable(s) with the greatest discrimination will be assessed next via quantification of consistency across studies.

RESULTS

At OMERACT 10, more than 107 participants attended the RA Flare Workshop, breakout groups, and final voting. In recognition of the importance of the patient perspective as a central and essential feature of developing clinically relevant outcome measures to assess flare, one of our patient research partners (JEM), with more than 40 years of person-

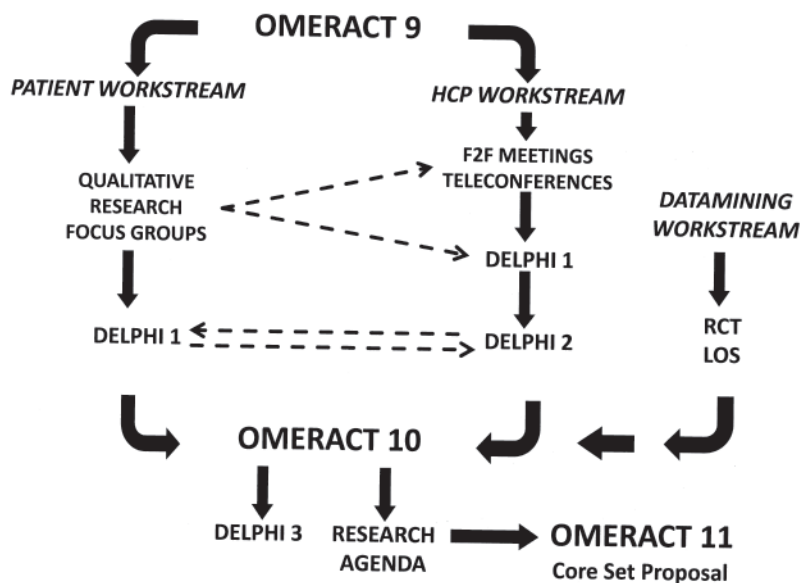


Figure 1. The coordinated RA Flare Working Group process and interactions. HCP: healthcare professionals; F2F: face-to-face; RCT: randomized clinical trials; LOS: longitudinal observational studies.

al experience of living with RA, presented his personal thoughts on the qualities of worsening disease activity, his expertise in flare self-management, and the impact of flares on his life. This was followed by presentations of data from the 4 work streams.

Data Presentations

Patient perspective. The patient perspective work stream was developed to acquire primary data from patients with RA regarding their experiences of disease flare/worsening. At OMERACT 9, initial data were acquired from the patient partners present that helped to inform a qualitative research approach to capture the experience of disease flare/worsening in RA patients. A series of 14 semistructured focus groups were conducted in the UK, USA, Canada, Australia, and Germany to explore the patient perspective¹⁰. Sixty-eight patients diagnosed with RA participated in these groups. All comments were transcribed (comments in German were translated into English, then back-translated into German), systematically analyzed, and coded independently by at least 2 researchers using inductive thematic analysis. All transcripts were reviewed by a patient research partner (JEM). A total of 289 codes, 28 categories, and 6 themes were identified. When data saturation was obtained, the themes were used as the substrate for subsequent Delphi exercises for patients and healthcare providers (discussed below).

In addition to information regarding the experience of flare, specific questions concerning prodromal or warning signs were addressed. Several important points emerged from the qualitative research. Patients identified 5 separate situations in which the term “flare” may be used: (1) a few bad days that would resolve over time on its own, (2) predictable and often self-induced worsening that could be attributed to over-activity, (3) worsening in a single bad joint, (4) predictable worsening due to external and emotional causes such as stress, and (5) worsening to the point that there was a need to speak with the rheumatologist about changing medication. There was also information on strategies of self-management. A detailed description of the process and comprehensive results from the qualitative research are reported in a separate publication¹⁰.

Delphi exercises. Separate preliminary Delphi exercises were conducted among HCP and patients in March and April 2010 via anonymous surveys completed on a secure site to identify additional domains that are needed in assessing flare and to indicate whether these were deemed essential, important, or not important to include in a definition of flare. Patients also had the option to complete a paper version of the survey. Each group was asked to identify and rank the 6 most important domains that were considered essential for a core set definition of “flare.” These Web-based activities were facilitated by IDENK, a UK-based consulting company engaged to facilitate the

work of different OMERACT groups. Patient-centered activities were overseen by appropriate Human Subjects Protection committees in the individual countries and institutions involved. The results of these Delphi exercises were presented at OMERACT 10.

Patient Delphi. Thirteen of the highest-ranking domains (encompassing 48 more detailed descriptors) from qualitative research were taken forward into a patient Delphi. These included joint symptoms, function, cognition, emotional distress, intimacy, sleep, fatigue, pain, self-management, and global experience/overall picture, participation, systemic features, and stiffness. The patient Delphi also captured data concerning 11 early warning signs identified in qualitative research. Preliminary results from 78 patients who had completed the Delphi before OMERACT 10 were presented. The patient Delphi did not include an assessment of the importance of laboratory features, or physician’s assessment, as these had not been raised by patients in the preceding qualitative research.

The top-ranking domains from patients were (percentage of patients indicating domain was essential or important): function (99%), fatigue (95%), participation (94%), systemic features (93%), pain (92%), stiffness (92%), and self-management (91%).

Data were obtained after the meeting from additional participants from the represented countries and a group of patients from France. Detailed results of the complete patient and healthcare professional Delphi exercises will be reported as a separate publication.

HCP Delphi. A 2-stage Delphi process was conducted for HCP to identify domains that may be important in assessing flare. An initial listing of domains used for Delphi 1 was obtained through literature review², face-to-face meetings, and teleconference, and expanded by the preliminary results of the patient Delphi. From 210 invitees, 100 completed the online Delphi 1, of whom 76% had OMERACT experience. The majority were physicians (82%), but there were also nurses, other HCP, and epidemiologists. Twenty-seven potential domains were included in Delphi 1, including the domains obtained from the patient qualitative research, with the inclusion of additional domains such as physician global assessment, laboratory values, and restrictions in participation, as defined by the ICF (International Classification of Functioning). There was also an opportunity for comments and additional domains to be included. In total, consensus (defined as > 70% agreement that a domain was essential or important) was achieved on 12 domains, which were carried forward to Delphi 2, the purpose for which was ranking of the identified domains to further clarify those considered essential. Items already considered as essential by > 70% of respondents from Delphi 1 (i.e., swollen joints, pain, tender joints, physical function, and patient global assessments) were included in the list of preliminary domains and not investigated further. Seventy-one of 102 invitees completed

the HCP Delphi 2, with similar characteristics by location and expertise. Delphi 2 was performed to evaluate a reduced list of 12 domains for which consensus was not achieved in Delphi 1 and to rank the identified domains from Delphi 1 that should be part of a flare core set, with a ranking of the top 6 domains from this list. The domains assessed in Delphi 2 were laboratory features, persistence of symptoms, fatigue, stiffness, participation, self-management, systemic features, extraarticular features, sleep/wake disturbance, symptom clustering, symptom unpredictability, and emotional distress. The top 6 domains (with the proportion of respondents listing as essential for inclusion in a core set) were laboratory features (83%), persistence (77%), fatigue (71%), stiffness (69%), participation (57%), and self-management (49%).

There was substantial agreement between HCP and patients for the importance of including existing core set domains in a RA flare definition (Table 1).

LOS evaluation. A subgroup of investigators in the data mining work stream have preliminarily evaluated data from UK and US LOS, to examine the performance of different components of RA symptoms and signs over time to detect patient worsening, and changes in these measures preceding a change or increase in therapy. These included the Consortium of Rheumatology Researchers in North America (CORRONA) database, the King's College Rheumatology Clinics Database, and the Brigham Rheumatoid Arthritis Sequential Study (BRASS) registry. Additional databases have been identified that could be queried for relevant information. In the King's College database, the flare anchor used was a worsening of physician or patient global assessment. SRM (defined as mean score change divided by the SD of the score change) were calculated for all clinical measures available according to the ACR Core Set for patients whose global data indicated "flare" versus those without flare, thereby introducing an operational group structure for patients with and without flare. In addition to physician global and patient global (the anchors), pain, and early morning stiffness score had moderate SRM (i.e., > 0.4). Disease Activity Score 28 (DAS28), swollen and tender joints, fatigue, and erythrocyte sedimentation rate exhibited small SRM (i.e., 0.2–0.4). Interestingly, the modified Health Assessment Questionnaire was not sig-

nificant, with SRM confidence intervals indicating no clear signal. In the BRASS study, a specific flare questionnaire was administered asking patients to recall and report information concerning flares over the preceding 6 months. Flares were reported between visits by 54%–75% of patients at 6 monthly visits over 3 years, with more than 30% reporting flares lasting more than 2 weeks; 43% reported short-lived flares lasting 3 days or less. SRM were calculated for clinical characteristics in patients reporting flare compared with those not reporting flare; Simplified Disease Activity Index (SDAI), Clinical Disease Activity Index (CDAI), patient global, physician global, pain, and HAQ were associated with SRM > 0.4. The results from the LOS flare data will be published as separate reports¹³. These studies together demonstrated that patient-reported flares and worsening of investigator and patient-reported disease activity are common and many episodes persist beyond 3 days. It was noted that the lack of a standardized method to capture disease worsening makes retrospective evaluation of existing datasets difficult. A coordinated effort will be needed to ensure that relevant domains are measured prospectively at the time of patient self-report of flare and that any approach that only periodically asks for patient self-report of numbers of flares between visits determines an appropriate recall period and methods to adequately assess flare duration and qualities. Future observational cohorts assessing disease worsening should include standardized questions inquiring about disease worsening to facilitate the examination of contributions of different domains that describe the concept of flare. These should be administered frequently in conjunction with collection of standard clinical and patient-reported variables to further this research agenda.

RCT evaluation. Several RCT in RA were identified to determine changes in core set measures that were associated with worsening RA disease activity. Commitments from 4 pharmaceutical company partners have been obtained to retrospectively analyze completed and reported RCT of biological agents in RA. These represent a range of compounds with different mechanisms of action.

Similar to the LOS, there was variability in the types of data collected in the RCT data sets. Data from RCT and LOS differ in that disease activity was high and improved with therapy in the RCT so that flares were less common than in the LOS studies. The initial analyses were based on a set of variables that were presumed to be relatively consistent from study to study with a plan to conduct a "meta-analysis" approach across studies⁷.

Analysis of one database was completed and presented at OMERACT 10. The analysis, based on an RCT database supplied by one of our pharmaceutical partners, provided preliminary sensitivity of various flare components or definitions available in RCT databases. The criterion measure used for a flare of disease activity was a worsening of physician global assessment of disease activity (i.e., change > 0).

Table 1. Agreement between healthcare professional (HCP) and preliminary patient Delphi exercises of existing core set domains.

Domain	HCP (%)	Patients (%)
Pain	100 (83)	92 (65)
Patient global	100 (82)	85 (43)
Function	100 (72)	99 (56)
Swollen joints	99 (88)	84 (53)
Tender joints	98 (76)	89 (53)
Physician global	94 (58)	Not asked

Subsequently, SRM were calculated for those subjects classified as having a “flare” versus those without (Figure 2).

Once again, composite measures including the SDAI, CDAI, and DAS28 were associated with moderate to large SRM (i.e., > 0.8); whereas tender joints, swollen joints, patient pain, and patient global assessment were associated with moderate SRM (0.5–0.6). Function, C-reactive protein, fatigue, and stiffness were associated with only small SRM, indicating low ability to discriminate between worsening and stable disease. Although based on preliminary data, this is in agreement with Vander Cruyssen, *et al*, who found that the momentary DAS28 score correlates best with the physician’s clinical judgment to change the medication in a cohort of RA patients under infliximab therapy¹⁴. A similar methodology is being used to evaluate additional RCT of other biological agents. Other criterion measures are also under consideration including changes in clinical variables in patients who had “flare” reported in adverse event data, withdrawal from study due to RA flare or disease worsening, and addition or increase in dose of medication (e.g., corticosteroids or NSAID) captured from concomitant medication data (to be reported in a separate publication).

Breakout Groups

After the plenary presentation of background and preliminary data, OMERACT 10 participants were assigned to

breakout groups for focused discussion on items for which there was either discordance between patient and HCP perspectives, and/or for which additional input was needed in developing the group’s research agenda. In more than half the groups, the facilitator or reporter was a patient research partner. In a reassembled plenary session, summary reports from the breakout groups were used to inform the plenary participants with an overview of the discussions and specific questions that were addressed in each group.

Consensus Voting

After presentation of the small-group summaries, the data presented were reviewed with the audience and a series of voting questions were presented to the group at large. It was agreed that there was no single clear indicator of flare that could be used across all datasets (91% yes, 3% no, 7% don’t know). Based on the widespread agreement between patient and HCP from parallel Delphi exercises, there was consensus (90% agreement) to include the following domains in a preliminary core set of measures for flare in RA: patient global assessment, pain, swollen joints, tender joints, function, and physician global assessment.

Additional voting was done in the final plenary session to achieve consensus on additional domains to evaluate for inclusion in a preliminary flare core set and research agenda (Table 2).

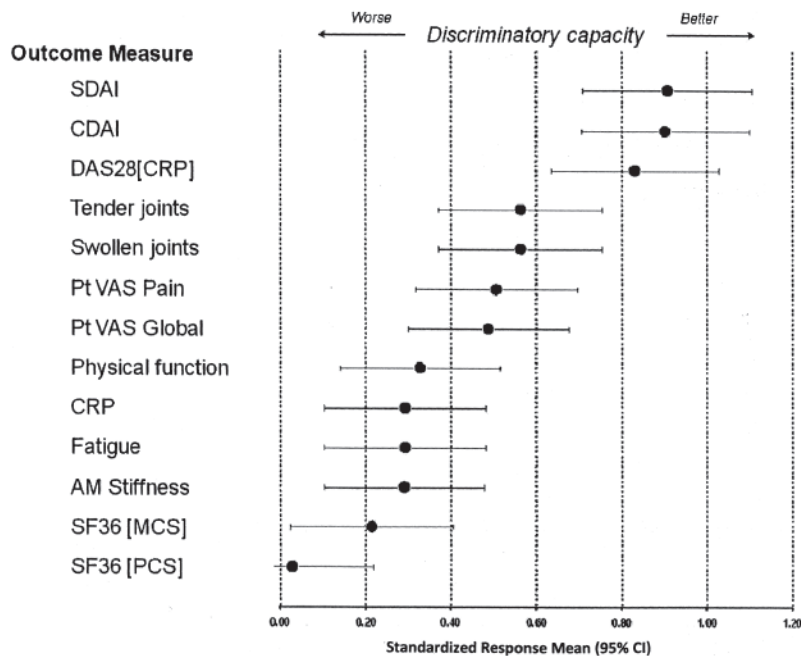


Figure 2. Standardized response means (SRM) for flare in a representative clinical trial. RA patients with an increase in physician global assessment from baseline to 6 months were defined as the “Flare” group (N = 192); the “No Flare” group (N = 262) was defined as those without an increase. SRM were calculated as the ratio of the group mean difference to the pooled SD of the mean change scores. SRM values with 95% confidence intervals show the discriminant capacity for flare in the available measures. MCS: mental component summary; PCS: physical component summary.

Table 2. Preliminary domains to evaluate flare in rheumatoid arthritis.

Patient global assessment*	Patient self-report of flare
Pain*	Fatigue
Swollen joints*	Stiffness
Tender joints*	Persistence
Physical global assessment*	Self-management
Function*	Participation
Laboratory features (e.g., ESR/CRP)*	Systemic features

* Existing RA Core Set measures. ESR: erythrocyte sedimentation rate; CRP: C-reactive protein.

There was consensus that a patient self-report of flare warranted further evaluation (92% yes, 4% no, 5% don't know). Consistent with findings from OMERACT 8¹⁵, there was consensus that fatigue should be included in a preliminary core set to measure flare (80% yes, 5% no, 15% don't know). In addition, stiffness also emerged as a domain to consider in a preliminary flare core set (70% yes, 18% no, 12% don't know). Persistence of symptoms (i.e., duration of worsening) was also recognized as important to take into account in assessing flare (75% yes, 12% no, 13% don't know).

There was consensus that laboratory features (75% yes, 13% no, 11% don't know) and systemic features (70% yes, 15% no, 15% don't know) should be included for further assessment in a flare research agenda. While consensus was not reached, more than 50% of participants voted that self-management (63% yes, 24% no, 14% don't know) and participation (56% yes, 34% no, 10% don't know) should be part of the ongoing research agenda.

Research Agenda

Based on the results from OMERACT 10, a research agenda was developed for OMERACT 11 in 2012. One of the first priorities was completion of a third Delphi round for patients and professionals combined, and for which participants of the earlier Delphi exercises and OMERACT 10 attendees will be engaged. It is hoped that this Delphi will include approximately equal representation of patients and HCP. An explanatory document that presented the major findings from OMERACT 10 was also needed to inform Delphi participants of relevant data. Based on the rankings obtained from Delphi 3, instruments will be selected for each of the domains that can be carried forward prospectively.

The consensus from OMERACT 10 also supported additional work to develop a tool that would capture patient self-report of flare, as well as a need to identify methods to accurately evaluate changes in stiffness, participation, and self-management. There was additional endorsement to evaluate the usefulness of the RA Impact of Disease (RAID) questionnaire to capture relevant aspects of worsening of disease activity¹⁶.

It was also agreed that additional interactions and discus-

sions between the OMERACT RA Flare Working Group and the OMERACT Remission in RA Working Group should continue. Although the voting through the first and second round of Delphi eliminated several domains from consideration in the preliminary core set (i.e., sleep/wake, emotional distress), there is a need to evaluate how to capture the key features of these domains to understand potential usefulness to detect flare.

As part of the research agenda, additional retrospective analysis of existing datasets and prospective incorporation of preliminary flare assessments will examine domains and instruments to fulfill the OMERACT filter of Truth, Discrimination, and Feasibility. Individuals from the OMERACT RA Flare Definition Working Group are working with the members of the working group defining core domains of the ICF in RA to ensure that domains selected can be appropriately mapped within the ICF context. Additionally, there is work under way to examine the use of the PROMIS framework, the NIH Patient Reported Outcomes Measurement Information System, to assist in capturing patient-reported measures that can be compared across diseases. The OMERACT RA Flare Definition Working Group will continue to seek additional global participants (e.g., from Latin America, Africa, and Asia) who can provide additional input into the activities.

It is anticipated that a significant portion of the research agenda will be accomplished for OMERACT 11 with a goal to conduct domain and instrument selection with analysis by the OMERACT filter in the context of RCT. Our ultimate goal is to develop a validated grading system that can be readily implemented across settings and activities to reliably quantify increases in disease activity (e.g., flares) from mild to moderate and severe. The research agenda will continue to examine flare in other contexts.

CONCLUSION

Patient partners in the OMERACT RA Flare Definition Working Group have consistently highlighted episodic RA worsening or flare in RA as a disabling, under-appreciated, yet integral feature of the overall RA experience. A well characterized and validated method to detect and measure RA flare is needed as a measure of response in RCT and to enhance timely communication between HCP and their patients regarding the impact of episodes of RA worsening/flares, and to enable treatment optimization and improve outcomes. An extensive research agenda was identified and cooperative efforts with other groups within OMERACT (e.g., Remission) and other organizations (e.g., ICF, PROMIS) were identified as desirable.

OMERACT 10 achieved consensus on the following domains in a definition of flare: patient global assessment, pain, swollen joints, tender joints, function, physician global assessment, and fatigue. There was also consensus that the following should be evaluated for inclusion in a flare

definition: patient self-report of flare, stiffness, persistence of symptoms, laboratory features, and systemic features. The results from OMERACT 10 provide preliminary information to facilitate RCT evaluating tapering of therapy. Recognition of flare and its impact and the features involved will enhance communication in understanding the time between visits or assessments in order to raise the bar for tight control.

ACKNOWLEDGMENT

We are grateful for the participation of patients and health professional participants in focus groups and Delphi exercises.

APPENDIX

List of study collaborators: OMERACT RA Flare Definition Working Group. Co-chairs: Clifton Bingham, Ernest Choy; Project Facilitator: Thasia Woodworth; Steering Committee: Rieke Alten, Clifton Bingham, Ernest Choy, Robin Christensen, Dan Furst, Sarah Hewlett, Amye Leong, James May, Thasia Woodworth; Fellows: Tessa Sanderson, Christoph Pohl. Group Members: Susan Bartlett, Christoph Pohl, Lyn Marsh, Vivian Bykerk, Jeff Curtis, Josef Smolen, Annelies Boonen, Mahboob Rahman, Niti Goel, Tim Shaw, Jim Witter, Swati Tole, Pam Montie, Kathleen Ferrell, Jasvinder Singh, Eswar Krishnan, Diane Moniz, Alan Solinger, Rick Brasington, Bruno Fautrel, David Magnusson.

REFERENCES

1. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S. Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. *Arthritis Care Res* 2010;62:647-56.
2. 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;36:2335-41.
3. van Vollenhoven RF, Boumpas D, Westhovens R, Brzosko M, Svensson K, Bjorneboe O, et al. Response, remission, and flare during early treatment of rheumatoid arthritis with infliximab: The REMARK Study [abstract]. *Ann Rheum Dis* 2010;69 Suppl3:535.
4. Den Broeder AA, Creemers MC, van Gestel AM, van Riel PL. Dose titration using the Disease Activity Score (DAS28) in rheumatoid arthritis patients treated with anti-TNF-alpha. *Rheumatology* 2002;41:638-42.
5. van den Bemt BJ, den Broeder AA, Snijders GF, Hekster YA, van Riel PL, Benraad B, et al. Sustained effect after lowering high-dose infliximab in patients with rheumatoid arthritis: a prospective dose titration study. *Ann Rheum Dis* 2008;67:1697-701.
6. Strand V, Crawford B. Improvement in health-related quality of life in patients with systemic lupus erythematosus following sustained reductions in anti-dsDNA antibodies. *Expert Rev Pharmacoeconomics Outcomes Res* 2005;5:317-26.
7. Bingham CO, Pohl C, Alten R, Christensen C, Choy EH, Hewlett SE, et al. "Flare" and disease worsening in rheumatoid arthritis: time for a definition. *Int J Adv Rheumatol* 2009;7:85-91.
8. Alten R, Pohl C, Choy EH, Christensen R, Furst DE, Hewlett S, 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;38:1745-50.
9. Woodworth T, Furst DE, Alten R, Bingham CO, Yocum D, Sloan V, et al. Standardizing assessment and reporting of adverse effects in rheumatology clinical trials II: the Rheumatology Common Toxicity Criteria v.2.0. *J Rheumatol* 2007;34:1401-14.
10. Hewlett S, Sanderson T, May J, Alten R, Bingham CO, Cross M, et al. "I feel like a tractor has run over me": Rheumatoid arthritis flare is more than a high joint count — an international patient perspective on flare where medical help is sought. *Rheumatology* 2011; [in press].
11. Hewlett S, Sanderson T, May J, Bingham CO, March L, Alten R, et al. "Just shoot me": RA patients experience flare as more than painful, swollen joints [abstract]. *Rheumatology* 2010;49 Suppl 1;141:285.
12. Hewlett S, Sanderson T, May J, Bingham CO, March L, Alten R, et al. Patients' words for scales to measure flare: "Sometimes it doesn't fit into that neat little box" [abstract]. *Rheumatology* 2010;49 Suppl 1;143:289.
13. Bykerk V. RA flares are common — BRASS [abstract]. *Arthritis Rheum* 2010;62 Suppl.
14. Vander Cruyssen B, Van Looy S, Wyns B, Westhovens R, Durez P, Van den Bosch F, et al. DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther* 2005;7:R1063-71.
15. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *J Rheumatol* 2007;34:1174-7.
16. Gossec L, Dougados M, Rinceval N, Balanescu A, Boumpas DT, Canadelo S, et al. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: a EULAR initiative. *Ann Rheum Dis* 2009;68:1680-5.