Establishing a Core Domain Set to Measure Rheumatoid Arthritis Flares: Report of the OMERACT 11 RA Flare Workshop

Vivian P. Bykerk, Elisabeth Lie, Susan J. Bartlett, Rieke Alten, Annelies Boonen, Robin Christensen, Daniel E. Furst, Sarah Hewlett, Amye L. Leong, Anne Lyddiatt, Lyn March, James E. May, Pam Montie, Ana-Maria Orbai, Christoph Pohl, Marieke Scholte Voshaar, Thasia Woodworth, Clifton O. Bingham III, and Ernest H. Choy

ABSTRACT. *Objective.* The OMERACT Rheumatoid Arthritis (RA) Flare Group (FG) is developing a data-driven, patient-inclusive, consensus-based RA flare definition for use in clinical trials, longterm observational studies, and clinical practice. At OMERACT 11, we sought endorsement of a proposed core domain set to measure RA flare.

Methods. Patient and healthcare professional (HCP) qualitative studies, focus groups, and literature review, followed by patient and HCP Delphi exercises including combined Delphi consensus at Outcome Measures in Rheumatology 10 (OMERACT 10), identified potential domains to measure flare. At OMERACT 11, breakout groups discussed key domains and instruments to measure them, and proposed a research agenda. Patients were active research partners in all focus groups and domain identification activities. Processes for domain selection and patient partner involvement were case studies for OMERACT Filter 2.0 methodology.

Results. A pre-meeting combined Delphi exercise for defining flare identified 9 domains as important (> 70% consensus from patients or HCP). Four new patient-reported domains beyond those included in the RA disease activity core set were proposed for inclusion (fatigue, participation, stiffness, and self-management). The RA FG developed preliminary flare questions (PFQ) to measure domains. In combined plenary voting sessions, OMERACT 11 attendees endorsed the proposed RA core set to measure flare with ≥ 78% consensus and the addition of 3 additional domains to the research agenda for OMERACT 12.

Conclusion. At OMERACT 11, a core domain set to measure RA flare was ratified and endorsed by attendees. Domain validation aligning with Filter 2.0 is ongoing in new randomized controlled clinical trials and longitudinal observational studies using existing and new instruments including a set of PFQ. (First Release March 1 2014; J Rheumatol 2014;41:799–809; doi:10.3899/jrheum.131252)

Key Indexing Terms:

OUTCOME AND PROCESS ASSESSMENT	FLARE	DISEASE ACTIVITY
OMERACT FILTER	CLINICAL TRIALS	RHEUMATOID ARTHRITIS

V.P. Bykerk, MD, Department of Rheumatology, Hospital for Special Surgery; E. Lie, MD, Department of Rheumatology, Diakonhjemmet Hospital; S.J. Bartlett, PhD, Johns Hopkins University, McGill University; R. Alten, MD, Schlosspark Klinik, Charité University Medicine; A. Boonen, MD, PhD, Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Center and Caphri Research Institute; R. Christensen, MSc, PhD, Musculoskeletal Statistics Unit, The Parker Institute, Department of Rheumatology, Copenhagen University Hospital; D.E. Furst, MD, David Geffen School of Medicine, University of California at Los Angeles; S. Hewlett, PhD, RN, University of the West of England, Bristol, UK; A.L. Leong, Healthy Motivation, Bone and Joint Decade; A. Lyddiatt, Musculoskeletal Group, Cochrane Collaboration; L. March, MD, PhD, University of Sydney Department of Rheumatology, Royal North Shore Hospital; J.E. May, National Fathers' Network; P. Montie, Arthritis Research Center; A-M. Orbai, MD, Johns Hopkins University; C. Pohl, MD, Schlosspark Klinik; M. Scholte Voshaar, Tools2use.eu; T. Woodworth, David Geffen School of Medicine, University of California at Los Angeles; C.O. Bingham III, MD, Johns Hopkins University; and E.H. Choy, MD, FRCP, Department of Rheumatology, Cardiff University.

Address correspondence to Dr. Bykerk, Department of Rheumatology, Hospital for Special Surgery, 535 East 70th St., New York, New York 07030, USA. E-mail: bykerkv@hss.edu

---

Flare in Rheumatoid Arthritis

The treatment goal of rheumatoid arthritis (RA) is remission or, when not possible, low disease activity[1,2,3]. The term "flare" is often used to describe episodes of disease worsening in RA, which are often unexplained and debilitating[4]. Flares can vary widely in frequency, duration, and intensity[5], and until recently, there has been little consensus on how to identify or define RA flares. For example, patients may use the word "flare" to describe any increase (transient or sustained) in RA symptoms such as joint pain, stiffness, or fatigue[6]. Conversely, in research and clinical care settings, healthcare professionals (HCP) often use the term to describe significant increases in inflammatory activity of sufficient intensity and duration that they necessitate the need for a change in treatment[4,7,8,9,10]. A reliable and valid method to identify and quantify significant inflammatory flares (hereafter referred to as flare) in severity and effect is needed given that biological drug tapering and withdrawal trials are now being conducted with "time to

flare" and "number of flares" proposed as outcomes. When treating individuals in clinical settings, it is important to determine whether prevention and/or rapid intervention to treat flares are needed to sustain remission, maximize productivity, and prevent joint damage and disability[11,12,13,14,15].

The Outcome Measures in Rheumatology (OMERACT) RA Flare Group (RA-FG), which includes researchers, clinicians, statisticians, and patients from Europe, North America, South America, and Australia, was established in 2006[7]. The goal of the group is to establish a validated outcome measure to identify RA flare[4]. This measure is intended to incorporate important elements/domains of flare from both the patient and HCP perspective. An international steering committee coordinates the efforts of the larger RA-FG, which comprises more than 40 members from academia and industry. Patient research partners (PRP) are integral members of this group, with a PRP representative serving on the steering committee.

Domains deemed to be important in measuring flare were identified by 14 focus groups conducted in Australia, Canada, Germany, UK, and USA comprising 67 patients unaffiliated with OMERACT[6] and from discussion groups of HCP and researchers who participated in special interest groups at OMERACT 8 and 9[11,13]. Next, separate Delphi exercises with a patient group and 2 iterative rounds with HCP were conducted and presented at OMERACT 10[4], after which a research agenda[4,10] was developed. The agenda included conducting a combined Delphi of attending patients and HCP[10], along with evaluation of instruments to assess the different clinical domains in clinical trials and longitudinal observational studies (LOS), and additional examination of some of the newly identified, patient-valued aspects of health related quality of life (Figure 1A). This development process aligns closely with newly developed OMERACT Filter 2.0 methodology (Figure 1B)[16].

We describe work conducted since OMERACT 10 and presented at OMERACT 11 resulting in endorsement of a core domain set to measure RA flare by OMERACT 11 participants.

METHODS

At OMERACT 11, the RA-FG presented a workshop with a goal to establish a core domain set to identify and measure flare in RA. During the workshop, background to the group and its leadup work from 2006 was outlined (Figure 1A). The role of patients as integral members of the RA-FG and its steering committee since the inception of this project was described by one of our PRP (AL)[17]. Results from the combined international patient and HCP stakeholder Delphi (Delphi 3) were presented[10], as well as preliminary data from LOS that examined feasibility, content, and known group validity of domains identified in the Delphi. Participants including patients, HCP, and researchers
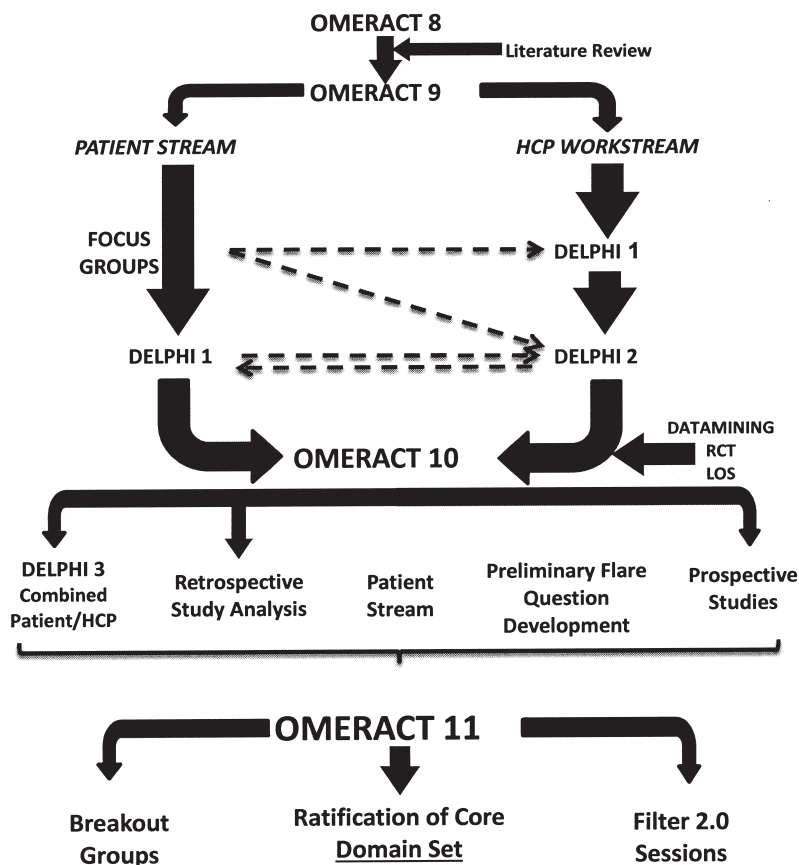
Figure 1A. Work plan of OMERACT RA Flare Working Group. HCP: healthcare professional; RCT: randomized controlled trials; LOS: longitudinal observational studies.
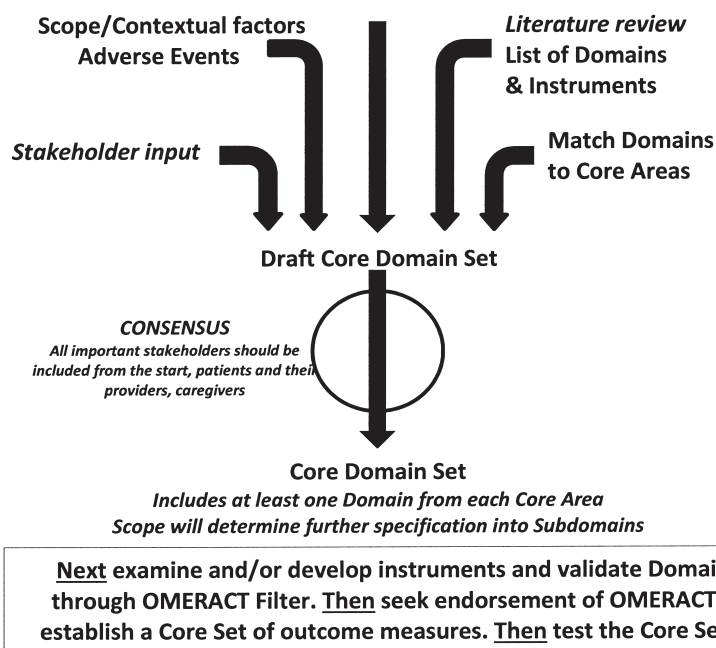

Figure 1B. Core domain set to measure RA flare aligns with OMERACT Filter 2.0 (from Boers M, *et al*. J Clin Epidemiol 2014; in press; adapted with permission[16]).

attended 1 of 5 breakout groups and, through in-depth discussion, they clarified issues on topics dealing with patient-HCP discordance, self-management, stiffness, and participation to help refine our research agenda. Throughout the meeting, discussions within the working group and among OMERACT delegates deepened understanding of the proposed domains in relation to flare and refinement of our research agenda. Questions were posed to the OMERACT 11 delegates for voting, including consensus for a core domain set to measure RA flare. A final plenary session vote was held on the last day.

Patients attending OMERACT 11 were active participants in all discussions, and some served as presenters, breakout group moderators, and as reporters. RA-FG work was also presented within the context of Filter 2.0 sessions as illustrative case studies for determining domain selection, including patient-reported outcomes (PRO) and engaging patients and PRP. The RA-FG work to develop a core domain set was demonstrated as aligning with Filter 2.0 methodology (represented in Figure 1B).

RESULTS
Presentation of Results of Delphi 3
After OMERACT 10, a combined Delphi exercise among more than 200 patients and HCP was conducted to identify core domains that should be measured for RA flare[10]. An *a priori* decision was made by the working group that domains would be considered "core" if > 70% of all participants rated a domain as essential to measuring flare. Several domains reached this combined 70% threshold (Pain, Function, Tender Joints, Swollen Joints, Stiffness, Patient Global Assessment, Participation, and Self-management; Figure 2)[10]. Other domains also achieved a 70% threshold by either patient or HCP consensus. Laboratory values [erythrocyte sedimentation rate (ESR)/C-reactive protein (CRP)] and evaluator global assessment were rated highly (75%, respectively) by HCP but not by patients; conversely, fatigue was endorsed highly (76%) by patients, but not by HCP. These data were presented during the initial plenary session with a proposal that (a) items reaching a delineation of 70% by both HCP and patients would represent central core domains to measure flare, (b) domains selected by > 70% of either group would be in an outer core, and (c) items that were rated as essential by 50–70% of either group would be placed on the research agenda. Subsequent discussions within breakout groups (below) and among delegates throughout the meeting resulted in a revised proposal that the RA core domain set for measuring flare should include domains rated > 70% by either patients or HCP. Domains identified by at least 50–70% of patients as being important in flare including sleep, systemic features, and emotional distress were retained in the research agenda as requiring additional evaluation (Figure 3). This revised framework was presented and ratified at the

final consensus voting session by the OMERACT delegates (discussed below).

Analysis of RA Flares in Longitudinal Observational Studies
The RA Flare Group has previously demonstrated that patient-reported flares are common events in clinical practice and captured in LOS[4]. To explore the feasibility, content validity, and known-group (discriminative) validity of the preliminary RA flare domains, data from 2 ongoing LOS were examined: the Norwegian Disease-modifying Antirheumatic Drug (NOR-DMARD) registry and the Canadian early arthritis cohort study (CATCH). We hypothesized that items representing each provisional flare domain would discriminate between patients experiencing a disease flare and those who were not in flare.

NOR-DMARD is an ongoing LOS of patients starting a new DMARD treatment with followup visits at 3, 6, and 12 months, and annually thereafter[18]. Patients who had initially responded to treatment were classified as being in flare using 3 different surrogate definitions: (a) patient-reported worsening reported as "worse or much worse" using a 5-point Likert scale; (b) increase in DMARD/systemic corticosteroid treatment; or (c) combined treatment change and patient-reported worsening. Analyses were conducted on changes between 3 and 6 months in variables including (a) domains within the existing RA core set of outcomes to measure disease activity, and (b) severity and duration of morning stiffness, fatigue, and participation (measured using the Medical Outcome Study Short Form-36 Survey domains). Briefly, results showed that scores on instruments commonly used to measure most of the proposed domains were significantly different between patients in flare versus those not in flare. Notably, morning stiffness severity discriminated better than morning stiffness duration. These data were presented by EL at the Fellow's oral presentation and as a poster during the meeting.

CATCH is a Canadian multicenter real-world LOS of patients with early RA, which began in 2006[19]. Patients are followed according to a standardized protocol: every 3 months within the first year and every 6 months thereafter. At study visits between November 2011 and April 2012, patients were asked whether they were in a disease flare at each visit; they also completed a set of preliminary flare questions representing domains contained within the RA-FG research agenda for flare as described below[20]. Analyses compared scores of patients who reported being in a flare versus those who did not. In brief, statistically signifi-cant differences (p ≤ 0.001) between flare and non-flare groups were observed for pain, fatigue, stiffness, function, participation, coping, patient global, and evaluator global assessments (all as measured on an 11-point numerical rating scale), and ESR/CRP.

These initial data provided preliminary evidence for the
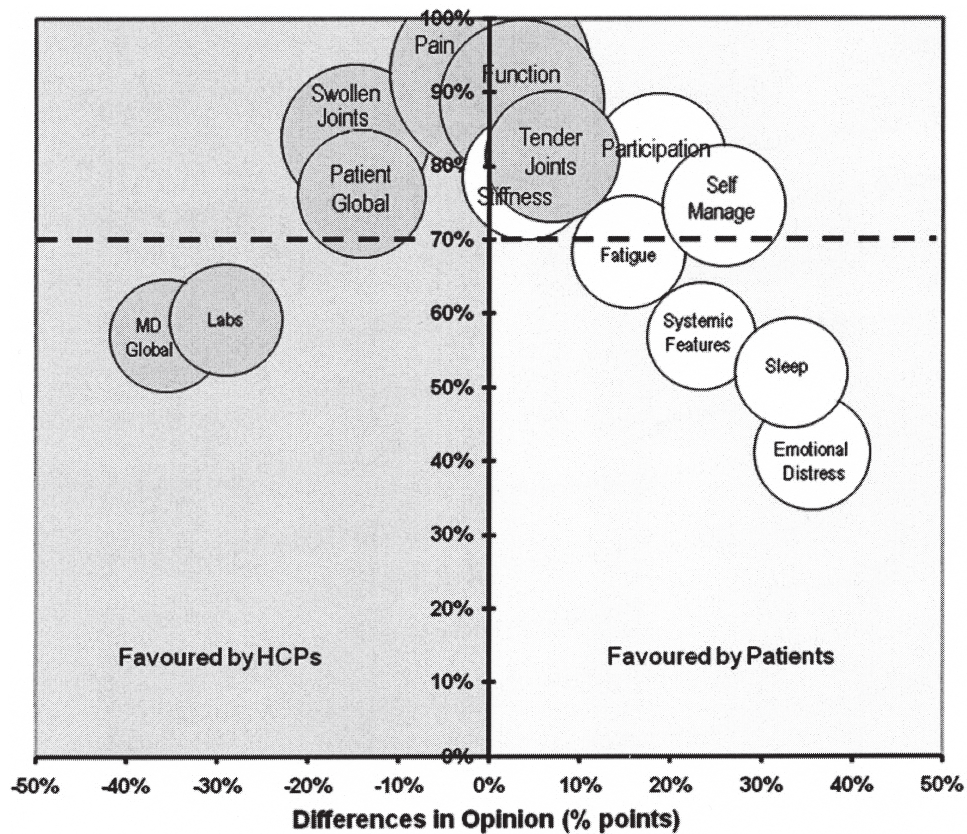
Figure 2. Consensus agreement for inclusion of domains used to define rheumatoid arthritis (RA) flare among patients and their healthcare professionals (HCP)[10]. Bubble size reflects the relative precision of each estimate. Domains reflecting the 1993 American College of Rheumatology preliminary core set for disease activity measures for RA clinical trials are shaded[33]. From Bartlett, et al. Ann Rheum Dis 2012;71:1855-60; used with permission.
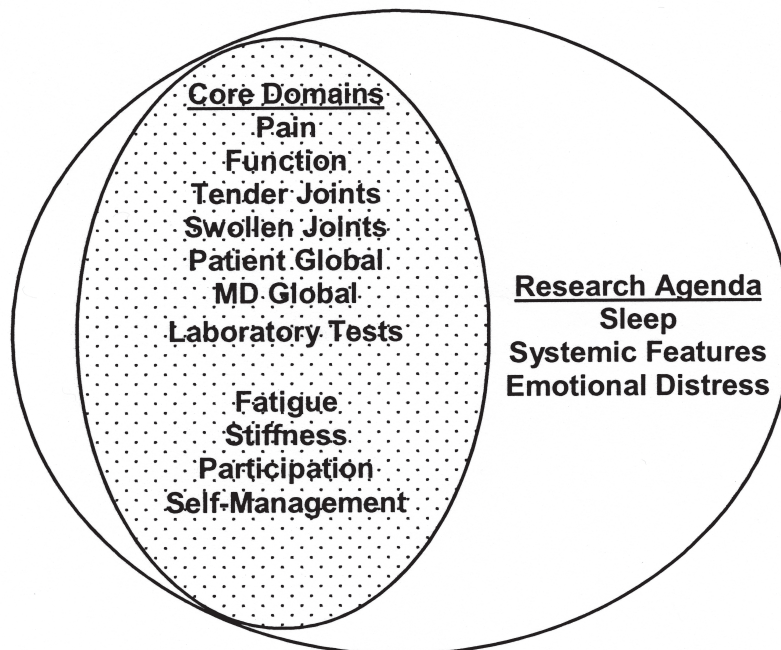
Figure 3. Core domain set to identify and measure rheumatoid arthritis flare.

feasibility and construct validity of the proposed RA core set domains to measure flare, and the ability of these proposed domains, as evaluated by a draft set of preliminary flare questions, to differentiate flare from non-flare (known-group validity). The full analyses concerning flare from NOR-DMARD have been published[21] and those from CATCH will be published separately (Bykerk, et al manuscript submitted 2014).

Patient involvement. Since the inception of the working group, the inclusion of patients' perspectives from patients at large and those collaborating as PRP has been an essential component of defining flare at all steps of domain development[7]. Some of the patients are representative members of the working group, and one (AL) is a member of the Steering Committee. RA-FG PRP actively contribute to the OMERACT Patient Perspective Working Group of OMERACT. Several PRP described their input, engagement, integration, and participation in the RA Flare Group process. A summary of the patient's experience and contributions was presented during the OMERACT 11 RA Flare Workshop[17] and in more detail in 2 Filter 2.0 sessions, a summary of which appears as an exemplar case study in Table 1[22].

Development of a Set of Preliminary Flare Questions to Explore Domains for the Core Set

As part of the research agenda at OMERACT 10, it was recognized that further information regarding the validity

Table 1. Research agenda for the Rheumatoid Arthritis Flare Working Group from OMERACT 11.

1. Identify existing instruments and develop new measures to evaluate each patient-valued domain included in the core set and research agenda for RA flare including:
   a. Self-Management (and evaluation beyond RA)
   b. Stiffness
   c. Participation
2. Examine other domains of importance to patients:
   a. Sleep
   b. Emotional distress
   c. Systemic Features
3. Assess and validate flare domains by analyzing OMERACT preliminary flare questions in several longitudinal observation studies and two randomized control trials
4. Employ methods, including item response therapy in the analysis and summation of multi-dimensional data for flare to evaluate thresholds of HCP and patient described domains associated with disease worsening
5. Assess domain measures in the context of a (i) "Patient global question" with well characterized and validated anchors, (ii) Patient assessment using a flare questionnaire in the context of available PROs at the time of each self-report, (iii) Physician/HCP assessments of disease activity, and (iv) treatment change.
6. Evaluate linguistic and cultural factors that pertain to domain identification and instrument selection and validation [Ref 32]
7. Determine whether flare should be represented as a change or a state (or both)
8. Determine duration/persistence criterion for defining flare

and responsiveness of the domains was needed[4]. In preparation for OMERACT 10, a scoping review of patient-reported outcomes to identify relevant questions and external anchors of flare for content and construct validation was conducted. A draft of potential items querying relevant flare domains was proposed by a PRP with subsequent modification involving additional PRP. The group also determined that it would be optimal to include (patient) ratings of severity and duration for all patient-defined flares. From earlier qualitative studies, patients[6] had identified self-management as an important consideration. At OMERACT 10, it was noted that the Rheumatoid Arthritis Impact of Disease (RAID) questionnaire[23] provided an expanded assessment of RA disease activity, and included some of the patient-valued domains relevant to flare; therefore, part of the research agenda was to consider RAID in assessing flare. Given this background, we developed a set of Preliminary Flare Questions (PFQ) that could be used in randomized controlled trials (RCT) and LOS to explore disease worsening from the patient's perspective using an 11-point numerical rating scale. These questions were constructed based on the combined experience of the group and built on the framework developed by RAID, using existing items from this and other validated questionnaires, while relying heavily on patients' words and descriptions obtained through our prior qualitative studies[6]. Questions were expanded with examples provided by PRP to ensure understanding of the range of symptoms and activities covered within a domain. These PFQ have been incorporated within several prospective LOS to acquire preliminary data on flare. These studies include CATCH, RETRO (REduction of Therapy in patients with RA in Ongoing remission), CARE (Cardiff Arthritis Real-time Evaluator), STPR (French Strategy of Treatment in patients with Rheumatoid Arthritis Flare Initiative)[24], and the Dutch Disease Activity Score 28-based flare studies. To provide validation information for the domains, an industry partner supported the translation and linguistic validation of the questions for use in an international RA clinical trial[25]. The questions have been translated into 13 languages (Spanish, German, Dutch, French, Portuguese, Danish, Hungarian, Italian, Polish, Romanian, Swedish, Catalan, and Russian), including linguistic adaptations for individual countries (e.g., German included adapted versions for Germany and Austria; Spanish included versions for Spain, Mexico, Colombia, etc.). Importantly, for each translation, there were 2 forward translations with reconciliation, followed by a bilingual clinical rheumatologist's review to ensure that intent and nuance were maintained. Next, there was back translation and pilot testing in 5 patients with RA from the country of interest. There was a final reconciliation of any changes for the certified version used.

Contemporaneously, with the work of the OMERACT group, a group of French investigators reported on their

work developing a measurement prototype to assess patient experience evaluations of recent RA flare events[24]. This group of investigators also evaluated many of the domains identified and presented by the RA-FG at OMERACT 10 and extended these findings with qualitative work from patients with RA in France, supporting the universality of the flare experience[8]. However, there are differences between the 2 prototype sets of questions. The former, built on patient report, was designed to identify flare retrospectively while the OMERACT PFQ could be used to assess current active flare episodes. To date, neither questionnaire has been validated in prospective clinical trials.

Breakout Groups

After the initial workshop plenary presentation, 73 delegates participated in 5 breakout groups on topics of self-management, patient-HCP discrepancy about defining a flare, self-management, stiffness, and participation. The breakout groups considered issues related to measuring proposed RA flare core domains and provided input into the research agenda and priorities for the workgroup. A PRP served as moderator or reporter in most groups and presented summary reports that informed meeting delegates of the discussion and issues addressed.

Self-management

Patient focus groups had revealed that patients self-manage flares using a number of strategies that include pacing, activity avoidance, increasing RA medications, and ultimately asking their HCP for help when symptoms remain severe and persistent, but usually only after the failure of self-management strategies[6]. In RAID and other studies, coping has been identified as an important part of the RA disease experience[23]. In breakout groups, discussants emphasized the importance of understanding the overlap and differences between coping and self-management. Most participants identified these as separate but interrelated constructs, because coping could also encompass other aspects of emotional health and social support as well as the use of self-management strategies. It was acknowledged that self-management often changes with disease duration as patients learn new strategies and adapt over time to living with RA. It was acknowledged that there are no instruments readily available to measure self-management and that this would be important to consider on the research agenda. In the CATCH study, self-management strategies varied considerably from doing nothing to doing many things[20].

Participation

"Activities and participation" is 1 of 2 components under functioning and disability of the International Classification of Functioning, Disability and Health (ICF) framework[26]. Our previous qualitative studies confirmed that participation in valued life activities is an important feature of health-related quality of life for people with RA[27]. For example, patients with RA reported their need to withdraw/isolate themselves and to ask for help at work and at home as an indicator used to assess the occurrence and severity of flare. Although some attendees suggested that measures of worker productivity also record elements of participation, all agreed that these measures would not adequately identify the varying effects of RA symptoms on participation in other activities (including social roles and leisure activities). There was agreement on the need to develop measures that could optimally assess participation across various settings.

In both the self-management and participation groups, discussants also noted that it could be important to assess these domains as both a "state" and as a "change", and it was acknowledged that these could be contextual factors. Both these domains remain challenging to define and measure. Existing instruments will need to be examined carefully and new instruments may need to be developed and validated.

Stiffness

During the Stiffness breakout session, data from recent RCT that incorporated stiffness as an endpoint were presented and discussed[28,29]. An American College of Rheumatology-OMERACT Fellow presented results of a scoping review, providing information on instruments, items, and anchors that have been used to evaluate stiffness. In addition, qualitative data from 2 focus groups exploring the patients' descriptions of stiffness in RA were presented. Data analysis consisted of an iterative consensus process among a team of clinical, social, and behavioral scientists. Overarching themes identified were pain-stiffness relationship, stiffness temporality, and alleviation-coping strategies. The OMERACT breakout group discussions noted that existing approaches that rely largely on estimations of morning stiffness duration were viewed as overly simplistic and do not address multiple aspects of stiffness. Additional data sets were identified by participants, in which stiffness (duration and severity) had been included as an outcome that could be potentially analyzed in the context of disease worsening. It was also recognized that experience of stiffness (local vs generalized) may be important to discern in RA and that stiffness is an important feature of other rheumatic conditions. There was a general agreement that stiffness was important to measure in RA flares and that further qualitative studies are needed to better understand this aspect of RA and other rheumatic diseases.

Consensus and Discordance between HCP and Patients in Identifying RA Flare

Delphi exercises and analyses of LOS have highlighted significant discordance between patients and HCP regarding

disease activity and the relevance of different domains in defining flare, although > 70% agreement exists for many domains. If occurrence of flare may be useful in guiding treatment decisions, understanding the reason for the discordance is important. Difference in attribution was suggested as a source of disagreement: patients may attribute changes in symptoms (e.g., worsening of pain) as central to a flare; whereas, HCP may focus on objective signs of increasing synovitis to recommend changing DMARD. OMERACT 11 participants agreed that it is important as part of the research agenda to explore how best to use patient-reported flare within the context of significant increases in inflammatory activity traditionally used by HCP to define significant inflammatory flares.

PRO Development

Workshops on PRO development noted the need to identify optimal contextual questions, addressing the potential for response shift and other forms of respondent bias and considering a potential weighting of domains as described for RAID[23]. Discrepancies between HCP and patients regarding which domains were deemed essential in identifying flare were viewed as evidence that both perspectives are important for identifying flares and may offer complementary information. OMERACT attendees highlighted the importance of additional work in translation and cross-cultural analysis when incorporating domain measurement instruments in a new flare PRO. Different approaches exist within item response theory methods (e.g., 1 vs > 1 variable models) that could place different emphases on the dimensionality of the flare experience and would be used in developing a new flare PRO. It was acknowledged that the US National Institutes of Health Patient Reported Outcomes Measurement Information System (PROMIS) also warranted evaluation as a measurement approach to evaluate RA symptoms related to flare. The research agenda for the group would thus include an assessment of the strengths and limitations of various assessment approaches.

Summary of Breakout Groups

Delegates continued to emphasize that inclusion of both PRO in combination with the traditional RA core set indicators is important in identifying and assessing RA flare. There was recognition that prospectively capturing changes in these characteristics within individuals was important, albeit challenging. Periodic assessments between visits to record symptoms and to decrease reporting bias using diary or digital media were suggested. It was acknowledged that more qualitative and quantitative research is needed to understand several key dimensions of flare including intensity, duration, and persistence of disease worsening[4,7]. Additional common themes included recognition that contextual factors such as setting (e.g., RCT vs LOS),

comorbidities, and disease duration and activity status also may be important. However, there was a common recognition that treatment for RA flares should be a shared decision between HCP and patients; moreover, while it was clear that patients and HCP may have different perspectives, these should be viewed as complementary rather than in opposition. Building consensus by recognizing similarities and differences, as recently discussed in an editorial, will be vital[8].

Consensus Voting

Initial vote (postplenary presentation). During the workshop, OMERACT delegates were presented with a number of questions upon which to vote. Participants overwhelmingly endorsed (97%) the importance of engaging both patients and HCP in identifying core domains, and 91% agreed that the RA Flare Group had sufficiently engaged patients and HCP throughout the process. Most agreed (85% agreed with 9% unsure and 6% disagreeing) that the methodology of using focus groups, literature review, expert and others' input, and iterative Delphi rounds was sufficient to identify core domains for RA flare, and incorporated all elements of the Filter 2.0 framework for establishing core domain sets (Figure 1B)[16]. In a preliminary vote immediately after the plenary session (but before breakout groups), 20% of respondents were unsure whether participation, stiffness, and self-management should be included in a core domain set to measure RA Flare. Feedback from delegates during the week indicated initial uncertainty about inclusion of these domains and may have reflected the large number of new attendees who were new to OMERACT, many of whom believed they did not have sufficient background information upon which to base their decision. Consequently, steering committee members engaged in discussion with delegates over the course of the week, debriefing information received from breakout groups, and conducting additional meetings as needed to identify gaps in understanding.

Final consensus voting. To address these gaps, additional information was presented to the OMERACT delegates, prior to the final vote at the end of the meeting. The proposed uses of a core domain set to measure RA flare and the process needed to fulfill the Filter 2.0 framework on core domain selection were described (Figure 1B)[16]. The initial criteria were revised to include, within the core, all domains deemed as important or essential by > 70% of either patients or HCP. A second and final round of consensus-voting to endorse a core domain set to measure RA flare was then conducted. There was 94% agreement that the existing RA core domains (pain, tender joints, swollen joints, physical function, patient global assessment, physician global assessment, and laboratory measures) with the addition of fatigue should be used. There was 91% agreement to include stiffness, 80% agreement to include participation, and 78%

agreement to include self-management as core domains to measure RA flare. Results are summarized in Figure 2.

DISCUSSION

The proceedings of the OMERACT RA Flare Group Workshop along with discussion of these in the context of Filter 2.0[22] have highlighted the importance of being able to identify and measure flare in RA. Through an iterative, data-driven process that included a wide range of relevant participants, including patients and PRP, the OMERACT RA Flare Group has now identified a core domain set to identify and measure RA flare for use in clinical studies. It is recognized that current methods of flare assessment and use of thresholds of existing composite measures may not be adequate or appropriate in many contexts[30]. Future OMERACT RA-FG will be able to incorporate Filter 2.0 methodology including incorporation of relevant contextual factors and mapping of the proposed core domain set to Filter 2.0 core areas, domains, and subdomains.

The process used by the OMERACT RA-FG has demonstrated the value of patient inclusion from the outset by PRP *and* by the broad group of patients who through various focus groups and Delphi consensus exercises were instrumental in identifying potential new core domains for RA flare. In addition, patients have enabled the RA-FG group to appreciate and conceptualize the role of participation (ultimate life effect measure) and self-management in defining RA flare.

The process of identifying these domains has been based on a bottom-up approach. Although initiated with a literature review, this methodology included focus groups of patients with RA from multiple countries, with careful qualitative analysis to define themes, and the input of relevant stakeholders. Through a 3-round Delphi process, patients and HCP identified domains that are important to measure when evaluating disease worsening in RA. Notably, 4 of these — fatigue, stiffness, self-management, and participation — were not part of the existing RA core set, although inclusion of fatigue had been strongly advocated[20,31]. The endorsement of this expanded set of core domains to measure flare represents an important development in the appreciation of patients' perspectives in measuring changes in disease activity[8].

While the results from this work have been grounded in a reference point of disease worsening, it seems likely that these domains are important potential indicators of disease activity throughout the course of disease regardless of the direction of change (improvement or worsening). This view is supported by the qualitative work of the RAID group that led to inclusion of assessments of fatigue and stiffness as well as coping in its instrument, in which coping may in some ways relate to self-management (or the failure thereof)[8,22]. This view is also supported by the ICF's emphasis, when describing and evaluating health, on the ability to participate in life activities[26]. The recognition of the conceptual underpinning of self-management as a means to judge the severity of symptoms is also a unique contribution of our work.

As yet there is incomplete agreement between patients and HCP as to the relative importance of each domain as it contributes to a measure of RA flare. In moving toward a flare "definition," it is possible that a patient-defined and a physician-defined flare will initially need to be recorded to understand the relative contribution of each domain, as it should be incorporated in a comprehensive measure of flare.

New challenges arise in evaluating flare using the core domain set to measure RA flare. Validated instruments already exist to measure some domains, but for domains participation, coping, and self-management new measures will need to be developed and tested; as for stiffness, further studies are needed to refine the relationship between stiffness and RA disease activity and to evaluate the validity of existing instruments.

The duration and intensity of flare may be important for making management decisions, particularly in the context of failure of self-management. Evaluating intensity with other features of RA disease worsening at a time when a patient self-reports a flare, along with the timing of subsequent reassessments to determine the duration and persistence of this flare[10], will be relevant. Thus further research as outlined in Table 1 will need to address current instruments and new measures to assess core and research domains, thresholds associated with worsening, and linguistic and cultural factors that pertain to domain identification and instrument selection and validation[32] to develop and validate criteria for flare.

At OMERACT 11, through workshops, breakout discussions, and plenary voting sessions there was consensus that 9 domains (pain, function, tender joints, swollen joints, patient global from the RA core set[33] with the addition of fatigue, stiffness, participation, and self-management as new domains) should constitute the core domain set for identifying and measuring RA flare. It was also agreed that additional domains (sleep, emotional distress, and systemic features) that met lower rates of consensus (> 50% but < 70%) would form part of the research agenda. Further research will evaluate situations in which there is discordance between HCP and patients in the identification of flare or worsening of disease activity. The inclusion of these domains may yield future insights into the pathophysiology of RA. Patient inclusion has been integral throughout the process of domain identification and incorporation, and PRP have been essential to this research. The incorporation of patient experience potentially offers new insights into RA and its management. Based on case discussion, this work aligns closely with the newly released OMERACT Filter 2.0 for Outcome Measures in Rheumatology[16].

REFERENCES

1. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010;69:631-7.

2. Smolen JS, Landewe R, Breedveld FC, Dougados M, Emery P, Gaujoux-Viala C, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. Ann Rheum Dis 2010;69:964-75.

3. Bykerk VP, Akhavan P, Hazlewood GS, Schieir O, Dooley A, Haraoui B, et al. Canadian Rheumatology Association recommendations for pharmacological management of rheumatoid arthritis with traditional and biologic disease-modifying antirheumatic drugs. J Rheumatol 2012;39:1559-82.

4. 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;38:1751-8.

5. Bykerk VP, Shadick N, Frits M, Bingham III CO, Jeffery I, Iannaccone C, et al. Flares in rheumatoid arthritis: frequency and management. A report from the BRASS Registry. J Rheumatol 2013 Dec 15; E-pub ahead of print.

6. Hewlett S, Sanderson T, May J, Alten R, Bingham CO, Cross M, et al. I'm hurting, I want to kill myself: rheumatoid arthritis flare is more than a high joint count; An international patient perspective on flare where medical help is sought. Rheumatology 2012; 51:69-76.

7. 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.

8. Bingham CO 3rd, Alten R, de Wit MP. The importance of patient participation in measuring rheumatoid arthritis flares. Ann Rheum Dis 2012;71:1107-9.

9. 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;38:1745-50.

10. Bartlett SJ, Hewlett S, Bingham CO 3rd, 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;71:1855-60.

11. Molenaar ET VA, Dinant HJ, Bezemer PD, Boers M, Dijkmans BA. Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. Arthritis Rheum 2004;50:36-42.

12. Lillegraven S, Prince FH, Shadick NA, Bykerk VP, Lu B, Frits ML, et al. Remission and radiographic outcome in rheumatoid arthritis: application of the 2011 ACR/EULAR remission criteria in an observational cohort. Ann Rheum Dis 2012;71:681-6.

13. Singh JA, Furst DE, Bharat A, Curtis JR, Kavanaugh AF, Kremer JM, et al. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res 2012;64:625-39.

14. Welsing PM, Landewe RB, van Riel PL, Boers M, van Gestel AM, van der Linden S, et al. The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis: a longitudinal analysis. Arthritis Rheum 2004;50:2082-93.

15. Schoels M, Knevel R, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas DT, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search. Ann Rheum Dis 2010;69:638-43.

16. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, D'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;in press.

17. Leong AMP, May J, Ferrell K, Voshaar M, Horgan B, Von Krause G, on behalf of OMERACT Flare Definition Work Group. Patient research partners shape participatory research progress to define rheumatoid arthritis flare: lessons learned to guide patient involvement [abstract]. Ann Rheum Dis 2001; Suppl:4449.

18. Kvien TK, Heiberg, Lie E, Kaufmann C, Mikkelsen K, Nordvag BY, et al. A Norwegian DMARD register: prescriptions of DMARDs and biological agents to patients with inflammatory rheumatic diseases. Clin Exp Rheumatol 2005;23 Suppl 39:S188-94.

19. Bykerk VP, Jamal S, Boire G, Hitchon CA, Haraoui B, Pope JE, et al. The Canadian Early Arthritis Cohort (CATCH): patients with new-onset synovitis meeting the 2010 ACR/EULAR classification criteria but not the 1987 ACR classification criteria present with less severe disease activity. J Rheumatol 2012;39:2071-80.

20. 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.

21. 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 2013.

22. Kirwan J, Boers M, Hewlett S, Beaton D, Bingham CI, Choy E, et al. Updating the OMERACT Filter: Core Areas as a basis for defining core outcome sets. J Rheumatol 2014;41: in press.

23. Gossec L, Paternotte S, Aanerud GJ, Balanescu A, Boumpas DT, Carmona L, et al. Finalisation and validation of the rheumatoid arthritis impact of disease score, a patient-derived composite measure of impact of rheumatoid arthritis: a EULAR initiative. Ann Rheum Dis 2011;70:935-42.

24. 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;71:1110-6.

25. Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 2005;8:94-104.

26. Reinhardt JD, Stucki G. Rheumatoid arthritis and participation—the social animal revisited. J Rheumatol 2007;34:1214-6.

27. 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.

28. Buttgereit F, Doering G, Schaeffler A, Witte S, Sierakowski S, Gromnica-Ihle E, et al. Efficacy of modified-release versus standard prednisone to reduce duration of morning stiffness of the joints in rheumatoid arthritis (CAPRA-1): a double-blind, randomised controlled trial. Lancet 2008;371:205-14.

29. Buttgereit F, Mehta D, Kirwan J, Szechinski J, Boers M, Alten RE, et al. Low-dose prednisone chronotherapy for rheumatoid arthritis: a randomised clinical trial (CAPRA-2). Ann Rheum Dis 2013;72:204-10.

30. Khan NA, Spencer HJ, Abda E, Aggarwal A, Alten R, Ancuta C, et al. Determinants of discordance in patients' and physicians' rating of rheumatoid arthritis disease activity. Arthritis Care Res 2012;64:206-14.

31. Aletaha D, Landewe R, Karonitsch T, Bathon J, Boers M, Bombardier C, et al. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. Ann Rheum Dis 2008;67:1360-4.

32. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000;25:3186-91.

33. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis Rheum 1993; 36:729-40.