ABSTRACT

Objective: To evaluate the reliability of concurrent flare identification using 3 methods (patient, rheumatologist and Disease Activity Score (DAS)28 criteria), and construct validity of candidate items representing the Outcome Measures in Rheumatology Clinical Trials (OMERACT) RA Flare Core Domain Set.

Methods: Candidate flare questions and legacy measures were administered at consecutive visits to Canadian Early Arthritis Cohort (CATCH) patients between November 2011 and November 2014. The American College of Rheumatology (ACR) core set indicators were recorded. Concordance to identify flares was assessed using the agreement coefficient. Construct validity of flare questions was examined: convergent (Spearman’s r); discriminant (mean differences between flaring/non-flaring patients); and consequential (proportions with prior treatment reductions and intended therapeutic change postflare).

Results: The 849 patients were 75% female, 81% white, 42% were in remission/low disease activity (R/LDA), and 16–32% were flaring at the second visit. Agreement of flare status was low–strong (κ’s 0.17–0.88) and inversely related to RA disease activity level. Flare domains correlated highly (r’s ≥0.70) with each other, patient global (r’s ≥0.66) and corresponding measures (r’s 0.49–0.92); and moderately highly with MD and patient-reported joint counts (r’s 0.29–0.62). When MD/patients agreed the patient was flaring, mean flare domain between-group differences were 2.1–3.0; 36% had treatment reductions prior to flare, with escalation planned in 61%.

Conclusions: Flares are common in rheumatoid arthritis (RA) and are often preceded by treatment reductions. Patient/MD/DAS agreement of flare status is highest in patients worsening from R/LDA. OMERACT RA flare questions can discriminate between patients with/without flare and have strong evidence of construct and consequential validity. Ongoing work will identify optimal scoring and cut points to identify RA flares.

Key messages

- Flares are common in rheumatoid arthritis and are often preceded by treatment reductions.
- Patients and MDs generally agree when patients flare, especially when previously in remission/low disease activity.
- OMERACT RA flare questions show evidence of reliability and construct, discriminant and consequential validity.

INTRODUCTION

People living with rheumatoid arthritis (RA) frequently experience transient increases in joint pain, swelling, and other symptoms such as stiffness and fatigue that indicate increased inflammation and worsening of their RA. Flares are generally expected to be reversible, though elevated RA disease activity persists in some cases.

Flares become clinically relevant when they are of sufficient intensity and duration to suggest that current therapy may be inadequate and a change in treatment may be required to optimise disease management. There is growing evidence of the importance of identifying flares in rheumatoid arthritis: reliability and construct validation of the OMERACT RA Flare Core Domain Set.

ORIGINAL ARTICLE

Identifying flares in rheumatoid arthritis: reliability and construct validation of the OMERACT RA Flare Core Domain Set

Vivian P Bykerk,1,2 Clifton O Bingham,3 Ernest H Choy,4 Daming Lin,2 Rieke Alten,5 Robin Christensen,6 Daniel E Furst,7,8 Sarah Hewlett,9 Amye Leong,10 Lyn March,11 Thasia Woodworth,12 Gilles Boire,13 Boulos Harauyi,14 Carol Hitchon,15 Shahin Jamali,16 Edward C Keystone,2 Janet Pope,17 Diane Tin,18 J Carter Thorne,18 Susan J Bartlett,3,19 on behalf of the OMERACT RA Flare Group and CATCH Investigators


ABSTRACT

Objective: To evaluate the reliability of concurrent flare identification using 3 methods (patient, rheumatologist and Disease Activity Score (DAS)28 criteria), and construct validity of candidate items representing the Outcome Measures in Rheumatology Clinical Trials (OMERACT) RA Flare Core Domain Set.

Methods: Candidate flare questions and legacy measures were administered at consecutive visits to Canadian Early Arthritis Cohort (CATCH) patients between November 2011 and November 2014. The American College of Rheumatology (ACR) core set indicators were recorded. Concordance to identify flares was assessed using the agreement coefficient. Construct validity of flare questions was examined: convergent (Spearman’s r); discriminant (mean differences between flaring/non-flaring patients); and consequential (proportions with prior treatment reductions and intended therapeutic change postflare).

Results: The 849 patients were 75% female, 81% white, 42% were in remission/low disease activity (R/LDA), and 16–32% were flaring at the second visit. Agreement of flare status was low–strong (κ’s 0.17–0.88) and inversely related to RA disease activity level. Flare domains correlated highly (r’s ≥0.70) with each other, patient global (r’s ≥0.66) and corresponding measures (r’s 0.49–0.92); and moderately highly with MD and patient-reported joint counts (r’s 0.29–0.62). When MD/patients agreed the patient was flaring, mean flare domain between-group differences were 2.1–3.0; 36% had treatment reductions prior to flare, with escalation planned in 61%.

Conclusions: Flares are common in rheumatoid arthritis and are often preceded by treatment reductions. Patient/MD/DAS agreement of flare status is highest in patients worsening from R/LDA. OMERACT RA flare questions can discriminate between patients with/without flare and have strong evidence of construct and consequential validity. Ongoing work will identify optimal scoring and cut points to identify RA flares.
of identifying and addressing inflammatory flare episodes as they may contribute substantially to worsening cardiovascular comorbidity, joint damage and other long-term outcomes. Although flares can occur unexpectedly, the risk of flare increases when RA treatments are tapered or withdrawn. In clinical trials, early identification of clinically important RA worsening would signal the need for (re)initiation of therapy. In clinical practice, early identification and resolution of flares would reduce the risks associated with persistently active disease to improve long-term outcomes.

Thus, there is a need for criteria and tools that can be used to reliably identify and quantify RA flares. Several methods have been proposed including a priori specified increases in disease activity scores (DASs).4,5 but there is little consensus. We are unaware of any reports that evaluate the reliability of flare identification which represents clinically important worsening by comparing different perspectives (eg, patients, treating rheumatologists, use of DAS worsening criteria). Similarly, a validated method to quantify flares remains elusive.

We have previously described our pathway to create a consensus-based definition of RA flares and identify the domains essential to include in any measure of flare. In brief, the Outcome Measures in Rheumatology Clinical Trials (OMERACT) RA Flare Group defined RA flares as episodes of increased RA disease activity accompanied by worsening symptoms, functional impacts, and clinical indicators of sufficient magnitude and duration to place individuals at greater risk of joint damage and poorer outcomes when left untreated. Our foundational qualitative and quantitative work with patients, clinicians and other scientists identified essential domains that could be used to measure flare severity.4,6 At OMERACT 2012, our RA Flare Core Domain Set was ratified by 200 + OMERACT attendees.7 The RA Flare Core Set included the American College of Rheumatology (ACR) RA core set8,9 (patient and physician assessment of disease, tender and swollen joints, acute-phase reactants, physical function, pain) and added fatigue, stiffness and participation. Exploration of self-management as a contextual factor10 and other research domains (systemic features, coping, sleep and emotional distress) was also endorsed.

The next steps needed to create a reliable tool to measure RA flare are to (1) identify candidate items to measure each RA flare domain; (2) evaluate the reliability of flare identification from different perspectives; and (3) assess the construct validity of our candidate flare items assessing each flare domain.11 To identify items, we reviewed conceptual models and existing patient-reported outcomes (PROs)12 and collaborated with members from the International Classification of Function and Health framework.13 Here, we present initial evidence of the reliability of flare identification and the construct validity of the candidate OMERACT RA flare items.

METHODS

Study participants

Data are from a subset of patients with RA seen over the first 2 years of follow-up in the Canadian Early Arthritis Cohort (CATCH) study, a prospective observational study of patients with early RA recruited at 19 centres across Canada initiated in January 2007.14 CATCH patients are followed every 3 months in year 1, and every 6 months in year 2 using a standardised protocol. Treatment generally follows Canadian guidelines for RA management.15 16 At each visit, patients complete validated measures of RA symptoms and function, the treating rheumatologist performs a physical examination to assess disease activity, and blood is drawn for analysis at local laboratories. Ethics boards at each centre approved the study, and written informed consent was obtained.

In November 2011, the candidate OMERACT flare questions were added to each visit. Here, we included all patients who met the 1987 ACR or 2010 ACR/European League Against Rheumatism (EULAR) criteria and had completed flare questions at ≥2 visits through November 2014. Data from the two most recent visits 3 months apart were used; when not available, data from the two most recent visits 6 months apart were selected (designated as V1 and V2).

Measures

Flare questions

Patients were asked whether their disease had worsened, remained the same or improved in the past week (7-point scale; much worse to much better) and if they were experiencing a flare of their RA (yes/no). Patients who classified themselves as flaring then rated the severity (11-point Numerical Rating Scale (NRS)) and duration (1–3, 4–7, 7–14 and >14 days) of their flare. All respondents then completed the candidate OMERACT RA Flare Core Domain Set items rating pain, physical function, fatigue, stiffness and participation over the past week due to RA using 11-point NRS (see online supplementary figure 1) and indicated tender and swollen joints on a 40-joint homunculus.20

Legacy PROs

Participants also completed the RAND-12,21 Rheumatoid Arthritis Disease Activity Index (RADAI)22 and Work Productivity and Activity Impairment-Rheumatoid Arthritis (WPAI-RA) Questionnaire.23 24

RA Indicators

ACR core set measures were recorded, including pain (10 cm Visual Analogue Scale (VAS)), physical function (Health Assessment Questionnaire-Disability Index (HAQ-DI)),24 patient and MD global assessments, MD tender and swollen joint counts (0–28), erythrocyte sedimentation rate (ESR) and C reactive protein (CRP); a DAS28 was calculated.
Classification of flare status reflecting clinically important worsening

Patient definition
Individuals who answered ‘yes’ to the question ‘Are you having a flare of your RA at this time?’ were classified as flaring.

MD definition
Physicians were asked: ‘Do you think your patient is in a flare today?’ (0 not in flare; 10 severe flare). Receiver operating curves were examined to establish the cut point which best reflected endorsement of flare (see online supplementary figures 2A–C).

DAS28 definition
As compared with the first of selected paired visits (V1), DAS28 worsening at the second visit (V2; ≥1.2 or ≥0.6 units if DAS28 at V1 was ≥3.2) was used to identify flares consistent with a definition used in recent studies.

Statistical analysis
Distributions were examined, and descriptive statistics were calculated. As data were missing for MD flare and DAS28 measures on some individuals, we used t tests to compare sociodemographic and RA characteristics between groups with and without complete data.

Reliability
Concordance between patient, MD and DAS28 identification of flares at V2 was assessed using the agreement coefficient. We hypothesised there would be moderate–high agreement (using Cohen’s criteria) of flare status between patients, MDs and DAS28.

Construct validation
To examine convergent validity, we used Spearman’s correlation coefficient to estimate the degree to which scores from flare domain questions correlated with each other, ‘legacy’ PROs measuring similar domains, and MD and patient-reported joint counts at V1. We hypothesised there would be moderate (r’s≥0.30) to high correlations (r’s≥0.50) between legacy PROs for flare questions assessing pain, function, participation, fatigue and stiffness; and weak (r’s≥0.10) to moderate correlations between pain and MD joint counts. We selected patients with good control of their RA at the V1 (ie, DAS28<3.2), and calculated the difference (95% CIs) in mean change scores at V2 in flare item scores, disease activity indicators and patient ratings of RA activity. We hypothesised that as compared with those not in flare at V2, flaring patients would have significantly higher scores on flare items and rate their RA as worse/much worse. Consequential validity was examined by evaluating the proportion of patients with treatment reductions at V1 and rheumatologist intention to escalate therapy at V2. Analyses were performed using SAS V9.3 (SAS Institute); p values <0.05 were considered statistically significant and all tests were two sided.

RESULTS
These analyses include data from 849 patients (see figure 1). The sociodemographic and RA characteristics of patients included in these analyses were similar to those who had completed <2 flare assessments (data not shown).

As shown in table 1, patients were mostly female, white and 57% were in remission or low disease activity (LDA). At V2, MD flare ratings were available for 85% (718/849), and DAS28 scores were available for 61%
(515/849 – the remaining 39% had missing ESR). Groups that were missing MD flare ratings or DAS28 scores did not differ from those with these data available on sociodemographic or RA characteristics, suggesting data could be treated as missing at random (see online supplementary table S1 for additional characteristics).

Reliability: agreement of flare status

At the second visit, 201/849 (24%) of patients classified themselves as flaring; MDs rated 233/718 (32%) and DAS classified 84/515 (16%) as flaring. Agreement of flare status varied based on disease activity level at the second visit. In patients previously in remission, agreement was high (κ’s=0.73) between patients and physicians and patients and DAS28 (table 2). For patients in LDA, agreement was moderate–strong between patients and physicians and patients and DAS (κ’s=0.44–0.63), but agreement was low for those in moderate–high disease activity (κ’s=0.17–0.35).

Construct validity of flare domains

At V2, OMERA CT RA flare domain questions correlated highly with each other (r’s 0.70–0.90; see online supplementary table S2) and with patient global (r’s 0.66–0.88), and moderately with patient joint counts (r’s 0.39–0.62; table 3).

Convergent validity

Across the three flare definitions used, correlations of OMERA CT RA flare domain questions with other PROs assessing similar constructs were high, with most >0.60 (see online supplementary table S2). Low–moderate correlations were observed between pain and MD/patient tender joint counts in flaring patients (r’s 0.29–0.48).

Known groups

At V2, mean scores for flare questions, other PROs and RA clinical indicators were significantly higher in patients who self-identified as flaring (table 4). Differences between groups were highest when patients and MD rated the patient as flaring. Patients who self-identified as flaring or also were rated by their doctors as flaring were much more likely (p<0.0001) to rate their RA as worse or much worse since the previous visit.

Consequential validity

Similarly, patient flare severity ratings and change in DAS28 scores at V2 were highest when patients and MDs agreed the patient was flaring (p’s=0.028 and <0.001,
Table 2  Agreement of flare status between patient, MD and DAS28 worsening criteria by disease activity level

<table>
<thead>
<tr>
<th>Flare classification (Visit 2)</th>
<th>Flare status</th>
<th>Patient</th>
<th>Agreement</th>
<th>MD flare classification (n=718)</th>
<th></th>
<th>DAS28 criterion flare classification (n=515)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Observed Expected (%)</td>
<td>AC1</td>
<td>≥1.2/0.6</td>
<td>&lt;1.2/0.6</td>
</tr>
<tr>
<td>All</td>
<td></td>
<td>98 (58%)</td>
<td>135 (25%)</td>
<td>512 (71%)</td>
<td>59</td>
<td>64%</td>
<td>31%</td>
</tr>
<tr>
<td>Remission</td>
<td></td>
<td>5 (23%)</td>
<td>39 (16%)</td>
<td>212 (79%)</td>
<td>78</td>
<td>66%</td>
<td>34%</td>
</tr>
<tr>
<td>Low disease activity</td>
<td></td>
<td>8 (50%)</td>
<td>17 (29%)</td>
<td>49 (66%)</td>
<td>59</td>
<td>55%</td>
<td>45%</td>
</tr>
<tr>
<td>Moderate disease activity</td>
<td></td>
<td>37 (69%)</td>
<td>40 (45%)</td>
<td>86 (60%)</td>
<td>49</td>
<td>57%</td>
<td>43%</td>
</tr>
<tr>
<td>High disease activity</td>
<td></td>
<td>23 (79%)</td>
<td>11 (65%)</td>
<td>29 (63%)</td>
<td>56</td>
<td>55%</td>
<td>45%</td>
</tr>
</tbody>
</table>

AC, agreement coefficient; DAS, Disease Activity Score.

Table 3  Relationship between items assessing OMERACT RA Flare domains and other measures by flare status

<table>
<thead>
<tr>
<th>Domains</th>
<th>N</th>
<th>Source</th>
<th>All</th>
<th>Patient</th>
<th>Physician</th>
<th>DAS28†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>849</td>
<td>201</td>
<td>485</td>
<td>84</td>
</tr>
<tr>
<td>Pain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How much pain due to RA in past week</td>
<td>91</td>
<td>0.91</td>
<td>0.89</td>
<td>0.89</td>
<td>0.91</td>
<td>0.90</td>
</tr>
<tr>
<td>Today’s level of pain today</td>
<td></td>
<td>RADAI</td>
<td>0.87</td>
<td>0.86</td>
<td>0.81</td>
<td>0.84</td>
</tr>
<tr>
<td>Pain past week</td>
<td></td>
<td></td>
<td>0.92</td>
<td>0.91</td>
<td>0.88</td>
<td>0.92</td>
</tr>
<tr>
<td>Joint area pain severity (0–48)</td>
<td></td>
<td>RADAI</td>
<td>0.71</td>
<td>0.55</td>
<td>0.67</td>
<td>0.58</td>
</tr>
<tr>
<td>Patient tender joint count (28)</td>
<td></td>
<td>Patient</td>
<td>0.62</td>
<td>0.36</td>
<td>0.57</td>
<td>0.48</td>
</tr>
<tr>
<td>MD tender joint count (28)</td>
<td></td>
<td>MD</td>
<td>0.49</td>
<td>0.29</td>
<td>0.42</td>
<td>0.36</td>
</tr>
<tr>
<td>Physical function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability score (0–3)</td>
<td></td>
<td>HAQ</td>
<td>0.76</td>
<td>0.73</td>
<td>0.68</td>
<td>0.74</td>
</tr>
<tr>
<td>Physical function</td>
<td></td>
<td>RAND-12</td>
<td>−0.65</td>
<td>−0.62</td>
<td>−0.58</td>
<td>−0.59</td>
</tr>
<tr>
<td>Daily activities in past 7 days</td>
<td></td>
<td>WPAI</td>
<td>0.77</td>
<td>0.67</td>
<td>0.74</td>
<td>0.75</td>
</tr>
<tr>
<td>Fatigue</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitality (RAND-12)</td>
<td></td>
<td>RAND-12</td>
<td>−0.63</td>
<td>−0.58</td>
<td>−0.59</td>
<td>−0.70</td>
</tr>
<tr>
<td>Unusual fatigue/tiredness past week</td>
<td>86</td>
<td>0.86</td>
<td>0.86</td>
<td>0.83</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>Participation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role—physical</td>
<td></td>
<td>RAND-12</td>
<td>−0.71</td>
<td>−0.69</td>
<td>−0.63</td>
<td>−0.67</td>
</tr>
<tr>
<td>Social function</td>
<td></td>
<td>RAND-12</td>
<td>−0.61</td>
<td>−0.61</td>
<td>−0.54</td>
<td>−0.67</td>
</tr>
<tr>
<td>Productivity while working</td>
<td></td>
<td>WPAI</td>
<td>0.78</td>
<td>0.73</td>
<td>0.70</td>
<td>0.77</td>
</tr>
<tr>
<td>RA affecting daily activities</td>
<td></td>
<td>WPAI</td>
<td>0.77</td>
<td>0.70</td>
<td>0.73</td>
<td>0.76</td>
</tr>
<tr>
<td>Stiffness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning Joint Stiffness Score</td>
<td></td>
<td>RADAI</td>
<td>0.69</td>
<td>0.51</td>
<td>0.66</td>
<td>0.57</td>
</tr>
</tbody>
</table>

*Spearman correlation coefficients for second visit.
†DAS28 scores <3.2 at second visit required an increase of 1.2 units whereas DAS≥3.2 at second visit required increase of 0.6 units to classify flare.
DAS, Disease Activity Score; HAQ, Health Assessment Questionnaire; OMERACT, Outcome Measures in Rheumatology Clinical Trials; RA, rheumatoid arthritis; RAND, Rheumatoid Arthritis Disease Activity Index; RAND-12, RAND 12 Health Survey; WPAI, Work Productivity and Activity Impairment.
respectively; see table 5). RA treatment had been reduced or stopped in 36% (10/28), and 61% of rheumatologists reported an intention to increase treatment in response to flare.

DISCUSSION
This is the first study to evaluate the reliability of concurrent flare identification in RA using three methods: patient reports, physician ratings and DAS28 criteria used in recent reduction/withdrawal trials. Agreement about flares using the three methods (patient, MD and DAS) was highest for patients initially in remission and LDA. Our data are also the first to show the construct validity of the five candidate OMERAICT RA flare domains (pain, fatigue, stiffness, function and participation). These data support additional evaluation of our patient-centered tool in treatment studies to establish numerical criteria for flare for use in trial and practice.

Agreement about flares varied and was influenced by disease activity level at the initial visit. Others have also reported that the patient and physician do not always agree about RA status, including flares.\(^5\)\(^\text{29–31}\) We found that agreement was highest between physician-based measures, either by MD report or DAS worsening, and patients, when disease activity had been in good control at the preceding visit. Conversely, agreement regarding flares was lowest between rheumatologists and DAS28 criteria when patients had previously been in moderate-high disease activity levels. This may, in part, reflect greater uncertainty in identifying worsening of disease when disease activity is already high. Also, in our study, the mean worsening in DAS28 of flaring patients was nearly two points higher, which is considerably greater than DAS criteria. It is not surprising that the least reliable method of identifying flare was the DAS criteria definition (identifying flare as ‘worsening’ of DAS28 that exceeded measurements error (0.6 units) or twice its value).\(^8\) This definition has been used in only a few studies that evaluated RA flares in the context of tapering or withdrawing trials.\(^6\)

Our results support the construct, discriminant and consequential validity of the five candidate items capturing the OMERAICT RA Flare Domain Core Set. Flare domain questions correlated highly with PROs measuring similar constructs that are widely used in RA trials and were able to discriminate between patients with/ without flare using three definitions. In people who were flaring at the second visit (but not the first), individual flare question scores were significantly higher, and similar increases were observed in other PROs as well as clinical indicators of RA disease activity. In contrast, there was little change in scores of non-flaring patients over these visits. Finally, in flaring patients, 61% of rheumatologists indicated they intended to intensify treatment indicating evidence of the consequential validity of flare identification and consistent with sufficient clinical worsening to justify escalation in therapy.
The five RA flare items represent patient-reported core domains identified by patients and providers as essential to measure for RA flare beyond the patient global assessment of disease activity. The high rates of agreement identifying flares between patients and physicians are not unexpected. The flare questions were developed to reflect the usual interchange of information between patients and physicians when discussing disease worsening. Concordance was lowest between patients and physicians when discussing disease status when patients were in moderate or high disease activity at the first visit. One reason for this may be that the DAS criteria do not directly incorporate any of the specific domains identified by patients and providers as essential to measuring RA flare, with the exception of patient global. While we found that patient global scores were highly correlated with the five RA Flare domain scores, it is unclear whether patient global alone can be used to identify flares related to inflammation. Notably, only two of the five domains, pain and physical function, are captured in the ACR RA core set, and core set measures make up standard composite measures of disease activity derived to measure improvement. RA composite measures alone may not be sensitive or specific to inflammatory flare, underscoring the need for a tool that can reliably identify and quantify RA flares.

Our results showed that agreement about flare status was high between patients and providers; also, the largest increases in flare domain scores and DAS28 occurred in flaring patients when there was patient/MD concordance if the patient was flaring. Better understanding of the factors related to discordance in flare reports between the patient and doctor assessments warrant additional study. In the interim, however, an important finding of our work for clinicians is recognition that patients can reliably identify significant increases in inflammatory activity, and that most of the time clinicians agree with patients when patients state they are in a flare. This is especially true in patients that had been in remission or LDA at a previous visit.

While doctors and patients may not always agree on flare status, agreement between both increases confidence that RA flares, which truly reflect worsening inflammation, can be reliably detected. A means to reliably identify and precisely quantify inflammatory flares in RA is needed for clinical trials where drug therapy is reduced or withdrawn as well as comparative effectiveness trials. In this large observational ‘real-world’ trial, rheumatologists classified 32% of patients as flaring, and treatment had been reduced/stopped at the prior visit in 36% of patients confirming that therapeutic change is an important antecedent to flare in many patients. As evidence grows that tight control is essential to improve long-term outcomes, early identification and treatment of flares seem essential to improve long-term outcomes. The ability to quickly and easily identify flares...
incorporating the patient perspective of flare could also facilitate more patient-centred care through the rapid identification of patients potentially requiring escalation of therapy between scheduled visits. Early flare identification enabled by patient report may also help patients know when to initiate additional self-management strategies.

Strengths of this study include the prospective collection of flare data in the context of a national observational study of early RA where many clinical and PROs were systematically collected in a standardised manner. Our sample included patients with all levels of disease activity. Foundational work closely followed OMERACT methods for developing new measurement tools and COSMIN criteria to ensure high methodological quality. However, there are limitations. We did not have a priori provide guidance or a definition of flare for patients and physicians, and the threshold identified by ROC curves to optimise discrimination of flare was low. Providing a standardised definition and querying physicians directly (yes/no) about flare status may increase agreement. Nevertheless, for patients initially under good control, agreement regarding flares was strong.

PRO measures are evolving, especially for fatigue, stiffness and participation. The ability to differentiate patients who are flaring or not was at the group level; additional work is needed to identify whether the questions can be used to reliably identify individuals who are flaring. Worsening assessed by patients and providers was at a single point in time; we did not specifically ask if judgements were made in relation to a previous time period (eg, 3 months), an approach being used by others. DAS28 criteria reflect changes from the previous visit, either 3 or 6 months prior. We have no information about symptoms, function, self-management (eg, transient use of glucocorticoids or non-steroidal anti-inflammatory drugs), and limited information on potential treatment changes between visits.

These initial results support the usefulness of the OMERACT RA flare questions for identifying flare. Additional evaluation is needed before they can be recommended for widespread use. Work is ongoing by our group to develop a scoring system for the OMERACT flare questions and to evaluate unidimensionality, responsiveness, and to identify clear thresholds that reflect worsening of RA inflammation signalling a potential need to intensify treatment. Flare data are being collected in international observational studies and randomised controlled trials (RCTs) of early and established RA to help establish criteria for symptom intensity and duration necessary to define inflammatory flare, and to develop thresholds for existing disease activity measures. Identifying and understanding the role of self-management strategies and other contextual factors also need to be considered. Further exploration of discordance between doctors and patients regarding RA flares and evaluation of patient and MD joint counts is also ongoing.

In conclusion, in routine rheumatology care settings, flares in RA representing clinically important worsening of inflammation are common and are often preceded by treatment reductions. Agreement about flare status among patients, treating rheumatologists, and using DAS28 criteria is high, especially for patients previously in remission or LDA. The five questions that represent the OMERACT RA Flare Core Domain Set, where patients are asked to rate their pain, fatigue, stiffness, function and participation have strong evidence of content validity, known groups and consequential validity. Additional work is also ongoing to develop a scoring system and identify the thresholds of change and flare severity that can be used by clinicians and researchers in order to reliably identify flares using the OMERACT RA flare questions.

Author affiliations
1Department of Rheumatology, Hospital for Special Surgery, Weill Cornell Medical College, New York, New York, USA
2Rebecca McDonald Center for Arthritis & Autoimmune Disease, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada
3Division of Rheumatology, Johns Hopkins School of Medicine, Baltimore, Maryland, USA
4Section of Rheumatology, Arthritis Research UK & Health and Care Research Wales CREATE Centre, Cardiff University, Cardiff, UK
5Schlosspark Klinik, Charité University Medicine, Berlin, Germany
6Musculoskeletal Statistics Unit, Department of Rheumatology, The Parker Institute, Copenhagen University Hospital, Bispebjerg and Frederiksberg, Denmark
7Division of Rheumatology, University of California, Los Angeles, Los Angeles, California, USA (Emeritus)
8University of Washington, Seattle Wash; University of Florence, Florence, Italy
9University of the West of England, Bristol, UK
10Bone and Joint Decade, Healthy Motivation, Santa Barbara, California, USA
11Department of Rheumatology, University of Sydney, Institute of Bone and Joint Research, Royal North Shore Hospital, St Leonards, New South Wales, Australia
12Division of Rheumatology, University of California, Los Angeles, Los Angeles, California, USA
13Division of Rheumatology, Université de Sherbrooke, Sherbrooke, Québèc, Canada
14Rheumatic Disease Unit, Institut de Rheumatologie, Montreal, Québec, Canada
15Arthritis Center, University of Manitoba, Winnipeg, Manitoba, Canada
16Vancouver Coastal Health Institute, Vancouver, British Columbia, Canada
17Division of Rheumatology, St. Joseph’s Health Care London, University of Western Ontario, London, Ontario, Canada
18Southlake Regional Health Centre, Newmarket, Ontario, Canada
19Divisions of Clinical Epidemiology, Rheumatology, and Respiratory Clinical Trials Unit, McGill University, Montreal, Québec, Canada

Acknowledgements The authors thank Laure Gossec, Maarten Boers, Vibeke Strand and the OMERACT Executive Committee for input which has shaped the direction of this work. The Canadian Early Arthritis Cohort (CATCH) investigators also include Majed Khaashi, Memorial University, St. John’s Newfoundland; Murray Baron and Ines Colmegna, McGill University, Montreal Quebec; Michel Zummers, HÔPITAL MAISONNEUVE ROSEMOUNT, Montreal, Quebec; Poonam Akravan, Lawrence Rubin, Birdee Kuriya, University of Toronto, Toronto, Ontario; Vandana Athilovia, Headwater’s Health Center, Orangeville, Ontario; William Bensen and Maggie Larche, McMaster University, Hamilton, Ontario; Lillian Barra, University of Western Ontario, London, Ontario; Bindu Nair, University of Saskatchewan, Manitoba; Christopher Penney, Dianne Mosher, Cheryl Barnabe, Glen Hazelwood, Calgary Health Sciences Center, University of Calgary, Alberta; Hector Arribilaga, Lethbridge, Alberta; Christopher Lyddiell, Grande Prairie, Alberta;
Rheumatoid arthritis

Alice Klinkhoff, University of British Columbia, Vancouver, Canada. They also thank Franci Sniderman for management of the CATCH project and Jim Wang, of McDougall Scientific for statistical support.


Contributors VPB, SJB and COB were involved in the conception, design, acquisition, analysis, interpretation, drafting and revisions of the manuscript. DL was responsible for the analysis, interpretation, drafting and revisions of the manuscript. EHC and LM were involved in the study conception, design, interpretation, drafting and revisions of the manuscript. RC, RC, DEF, SH, AL, LM, TW and KV were responsible for the conception, design, interpretation and revisions of the manuscript. VB, JP, GB, CH, SJ, DT, JCT and ECK participated in the conception, acquisition, interpretation and revisions of the manuscript. All authors reviewed and approved the final manuscript.

Competing interests OMERACT is an international rheumatology outcomes methodology group that has received hands-off funding from more than 23 pharmaceutical and clinical research companies over the last 2 years. COB and LM are members of the OMERACT Executive Committee but receive no financial remuneration for their service in this role. We thank UCB, Inc. for providing funding to support translation of the PFI into 13 languages (Spanish, German, Dutch, French, Portuguese, Danish, Hungarian, Italian, Polish, Romanian, Swedish, Catalan and Russian), including linguistic adaptations for individual countries (eg, French included versions France and Canada) using validated methods for use in an international RA clinical trial. Additional unrestricted funding for the OMERACT RA Flare Group was provided by Pfizer (Germany), Novartis and Actelion. The CATCH study was designed and implemented by the investigators and financially supported initially by Amgen Canada Inc. and Pfizer Canada Inc. via an unrestricted research grant since the inception of CATCH. As of 2011, further support was provided by Hoffmann-LaRoche Ltd., UCB Canada Inc., Bristol-Myers Squibb Canada Co., AbbVie Corporation (formerly Abbott Laboratories Ltd.), Medexus Inc. and Janssen Biotech Inc. (a wholly owned subsidiary of Johnson & Johnson Inc.). COB and SJB were supported in part by a Patient-Centered Outcomes Research Institute (PCORI) Pilot Project Award (1U2P-H000751-01) and a PCORI Improving Methods for Conducting PCOR Award (SC14-1402-10818). All statements in this report, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of PCORI, its Board of Governors or Methodology Committee. VPB is supported by the Cedar Hill Foundation, New York, NY and by NIH grant (1U1HR0067691). RC The Parker Institute is supported by grants from the Oak Foundation.

Ethics approval Central Ethics Committee at each institution where study was performed.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES


37. Hewlett S, Dures E, Almeida C. Measures of fatigue: Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire (BRAF MDQ), Bristol Rheumatoid Arthritis Fatigue Numerical Rating Scales (BRAF NRS) for severity, effect, and coping, Chalder Fatigue Questionnaire (CFQ), Checklist Individual Strength (CIS20R and CIS8R), Fatigue Severity Scale (FSS), Functional Assessment Chronic Illness Therapy (Fatigue) (FACIT-F), Multi-Dimensional Assessment of Fatigue (MAF), Multi-Dimensional Fatigue Inventory (MFI), Pediatric Quality Of Life (PedsQL) Multi-Dimensional Fatigue Scale, Profile of Fatigue (Prof), Short Form 36 Vitality Subscale (SF-36 VT), and Visual Analog Scales (VAS). *Arthritis Care Res (Hoboken)* 2011;63(Suppl 11):S263–86.
Identifying flares in rheumatoid arthritis: reliability and construct validation of the OMERACT RA Flare Core Domain Set


*RMD Open* 2016 2:
doi: 10.1136/rmdopen-2015-000225

Updated information and services can be found at: http://rmdopen.bmj.com/content/2/1/e000225

*These include:*

**Open Access**
This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to: http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to: http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to: http://group.bmj.com/subscribe/