"Flare" and Disease Worsening in Rheumatoid Arthritis: Time for a Definition

Clifton O Bingham III, MD[1], Christoph Pohl, MD[2], Rieke Alten, MD[2], Robin Christensen, MSc, PhD[3], Ernest H Choy, MD, FRCP[4], Sarah E Hewlett, PhD, RN[5], James E May, MA[6], Vibeke Strand, MD[7], Thasia G Woodworth, MD[8], and Dan E Furst, MD[9]

[1]Johns Hopkins University, Baltimore, MD, USA; [2]Schlosspark Klinik, Charité University Medicine, Berlin, Germany; [3]The Parker Institute: Musculoskeletal Statistics Unit, Frederiksberg University Hospital, Copenhagen, Denmark; [4]Academic Department of Rheumatology, King's College London, London, UK; [5]University of the West of England, Bristol, UK; [6]Seattle, WA, USA; [7]Stanford University, Portola Valley, CA, USA; [8]Roche Products, Welwyn Garden City, UK; and [9]UCLA, Los Angeles, CA, USA

Submit comments or questions for the authors at
www.advancesinrheumatology.com

"Flare" is a term that is commonly used by both patients and clinicians to describe a debilitating worsening of symptoms. While conceptually recognized, there is limited formative research to identify the measures of clinical characteristics or other variables that define this state. As treatment strategies and medications become available to achieve low levels of disease activity and remission, measures to guide tapering of medications, or re-treatment when disease appears to worsen, are needed to facilitate clinical trials and to move these treatment approaches from clinical trials to clinical practice. While several different definitions of flare or disease worsening have been used in clinical studies, the validity of these to guide changes in treatment has not been tested. Moreover, the experience of RA patients has not been adequately evaluated to assure understanding of features of disease worsening that are most important, their magnitude, and/or their temporal occurrence and duration, which then cause a patient to seek change in treatment. An Outcome Measures in Rheumatology Clinical Trials (OMERACT) initiative, which includes researchers, clinicians, and patients, is now underway to evaluate these aspects of disease worsening in RA. Timely, data-driven completion of this project with an agreed definition of RA flare is intended to facilitate clinical trials, observational clinical research, and ultimately enhance clinical care. *Int J Adv Rheumatol* 2009;**7**(3):85–91.

It has long been recognized that rheumatoid arthritis (RA), even when well treated, may have periods of exacerbation and improvement. Both patients and their clinicians have called these periods of worsening an RA "flare". Some of these exacerbations are short-lived, necessitating minimal or no intervention, while others may be of a sufficient severity to require changes in medications that address the symptoms and loss of function that accompany a flare. While this phenomenon is well recognized by both physicians and patients, there are no generally agreed-upon parameters to define disease worsening, to characterize its severity, or to describe its onset and duration.

RA patient care has undergone important change in the last two decades, with the ability to diagnose disease earlier, the recognition of early joint damage as leading to disability, and the ensuing advent of effective treatment paradigms often including biological therapies. With the increasingly common use of aggressive and earlier combination disease-modifying antirheumatic drug (DMARD) treatment, the numbers of patients experiencing substantial improvement – either

Address for correspondence: Clifton O Bingham III, MD, Division of Rheumatology, Johns Hopkins University, 5200 Eastern Avenue, Mason F Lord Center Tower Room 404, Baltimore, MD 21224, USA. Email: clifton.bingham@jhmi.edu

Table 1. Examples of the use of a definition of RA flare.

Usefulness in Clinical Trials
• Evaluating duration of effect • Evaluating changes in drug doses, frequencies, or strategies for tapering • Evaluating the need for dose escalation or re-administration Evaluating the efficacy of concomitant medications (e.g. steroids, NSAIDs) to treat increased or oscillating symptoms • Capturing and quantifying certain "adverse events" (e.g. "RA Flare", lack of efficacy withdrawals)
Usefulness in Clinical Practice
• Deciding when to increase dose or add or change drugs based on uniform definition of disease worsening (e.g. RA "flare") • Developing guidance to re-administer periodically dosed medication • Defining what is an "acceptable" degree of disease oscillation, amount, and duration, incorporating the patient's perspective • Detecting the earliest or prodromal signs and symptoms of impending "flare" to potentially avert severe disease worsening • Improving evaluation and communication of patient reported outcomes and self-management strategies to limit the occurrence of disabling disease worsening

NSAIDs: nonsteroidal anti-inflammatory drugs; RA: rheumatoid arthritis.

remission or very low disease activity – is increasing. Recent clinical trials are reporting data for "remissions" by various criteria (e.g. 28-joint count Disease Activity Score [DAS28], Clinical Disease Activity Index [CDAI], and Simplified Disease Activity Index [SDAI]), as well as high magnitude levels of improvement from the initial treatment baseline, such as the 90% improvement in American College of Rheumatology criteria for RA (ACR90) measure [1]. The Outcome Measures in Rheumatology Clinical Trials (OMERACT) working group have an ongoing effort to develop a definition for a minimal disease activity state [2,3], and the American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) are working to develop a revised definition of remission that can be used for both clinical trials and clinical practice [4]. With such outcomes now possible, patients and clinicians are increasingly interested in guidelines for tapering medications to maintain a low disease activity state (**Table 1**). However, developing study designs to test or compare such strategies is limited because there is no standard definition for disease worsening (often characterized as "flare"). The development of a standard definition could also enable clinical treatment decisions, using objective criteria to evaluate RA "flares" for proactive control of disease and enhanced patient management.

Importance of defining flares in clinical practice

Clinicians caring for patients with RA appreciate the temporal fluctuations of disease activity, which may vary in magnitude and impact for the patient. Even patients who are well controlled describe "good days" and "bad days". The term flare has entered the vernacular to describe various levels of disease worsening by both patients and physicians, but there is no commonly accepted understanding of the constituents of a flare. Some clinicians believe this to be an increase in joint swelling and tenderness, while others follow a rise in acute phase reactants to detect a "flare"; others use characteristics of disease such as patient or physician global assessments to "document" worsening. Confusing this issue further, it may be that worsening in a single joint as well as a mild worsening in multiple joints could be interpreted as a flare. Thus, there is heterogeneity in what may constitute a flare of disease, but more importantly, the decision to act upon these factors in terms of specific interventions also varies widely from patient to patient and practitioner to practitioner. Finally, a goal of effective treatment should be to decrease or eliminate flares or the frequency of periods of disease worsening, as these may result in a progressive loss of function and time lost from regular daily activities, including work.

While some might view decision-making based on individual characteristics as part of the "art" of medicine, approaches vary widely, and there are no best practices. Because many patients are not cared for by rheumatologists on an ongoing basis, the provision of some guidance for practitioners and patients may be helpful to set thresholds for contacting the rheumatologist for intervention. There is also an important aspect of understanding flares in all their different perspectives, to assist in management guidelines for specific therapies.

Interventions used to treat a flare range from adding, switching, or increasing the dose of medications including nonsteroidal anti-inflammatory drugs (NSAIDs), corticosteroids, DMARDs, and biologics. If a patient comes for a visit with a mild degree of worsening in disease activity, it must be determined whether such a change represents part of an expected oscillation of RA disease activity that could be managed with non-pharmacological intervention, or whether the magnitude or duration of worsening mandates a more substantial intervention. Additional considerations when evaluating a flare are the severity of a symptom or symptom complex, and the persistence of worsening rather than simply an assessment at a single point in time. Changes in treatment regimens, in particular biologic regimens, may substantially increase cost (and possibly risk) due to premature repeat administration, early dose escalation, or changes in frequency of dosing. Thus, it becomes increasingly important to establish both a conceptual and operational framework to provide guidance for such decisions.

Over the last decade, substantial improvements in outcomes in patients with RA, both as a result of earlier DMARD intervention and due to the introduction of more potent regimens of combinations of DMARDs, including biological agents, have been seen. Detectable differences, while once an acceptable goal of therapy, have been replaced by an

expectation for significant, clinically meaningful improvements of disease, initially described for clinical trials but now clearly playing out in clinical practice. Now, the goals of therapy are to achieve a state of remission or minimally active disease, and these endpoints – once used only for clinical trials – are now being applied to clinical practice (e.g. DAS28 remission) [5–8].

"There is no commonly accepted understanding of the constituents of a flare"

With the ability to induce low activity states, practitioners and patients ask the questions of "how much?" and "for how long?" Once an acceptable degree of disease control is achieved, many patients begin to self-taper medications, and physicians also begin to modulate therapy. Yet these decisions are largely guided by individual factors, and there is a very limited literature to provide guidance for tapering medications. Furthermore, the decision of which drug to taper is variable and additional questions continue to emerge: does one initially taper corticosteroids, taper methotrexate, remove other combination DMARDs, or taper biologic dose or frequency? If such a change in therapy is introduced, how long should one wait before beginning the next step down? Which outcome variables are most appropriate to determine how much worsening is allowable before stepping therapy back up?

There is also limited information or recommendations regarding repeat administration of periodically administered medications with a variable infusion schedule, such as rituximab. Recently, consensus-based recommendations have been provided regarding retreatment with rituximab for patients in DAS28 >0.6 "or an equivalent change in disease activity", or residual disease defined by DAS28 as high activity (>5.1) [9]. While such consensus-based definitions may be helpful, the validity of such statements needs to be tested in both clinical trials databases as well as in clinical practice cohorts.

How can insights from patients help?

The above discussion has largely focused on the physician–practitioner perspectives of disease worsening. Patient-reported outcomes are increasingly recognized as a critical component of patient assessment in clinical trials and clinical practice. Minimal detectable differences (MDD) and minimal clinically important differences (MCID) in terms of global assessments, pain, the Health Assessment Questionnaire (HAQ), and health-related quality of life (HRQoL) by Short Form-36 (SF-36) have been established for improvement and for worsening. These established differences indicate that patients perceive worsening sooner, with less change than for improvement in each measure. These instruments also reflect the effect of confounding factors such as duration of disease and number of prior DMARDs. Whereas the HAQ may show less improvement when more impairment in physical function is present, other instruments such as SF-36 may show more change in those domains with the lowest scores at baseline [10]. Patients perceive and report their HRQoL relative to severity of disease at start of treatment and magnitude of improvement, and possibly other factors. Thus a patient's tolerance for a small degree of worsening may differ given a longer term understanding that small fluctuations may be part and parcel to the disease experience [11].

The involvement of patients has been an imperative part of the OMERACT process. Patient representatives have had an important role in helping professionals to understand the totality of the disease experience. The importance of this perspective has been seen in the recognition that fatigue is a critical domain in RA that was not adequately captured in many historical cohorts or clinical trials [12]. From a patient-focused perspective, many new questions emerge regarding worsening of disease. These are poorly studied and include the following:

- What is a "flare"?
- Does this term capture the experience of worsening?
- Are there different types of flares?
- How long must symptoms persist and how severe must symptoms become until worsening requires intervention?
- What types of self-management strategies are used?
- How effective are different medical interventions in alleviating symptoms?
- What is the anchoring point for judging worsening?
- Are thresholds of detection different for improvement and worsening?

The OMERACT "flare" group is conducting a rigorous qualitative research study to explore the meaning of flare from the patient perspective, comprising 15 focus groups of RA patients. Although data collection and analysis are still ongoing, it is already clear that flare has many different meanings and contexts for individuals, and that our existing, rather loose, terminology may not capture the different nuances important to patients. Emerging data have yet to be confirmed in the full data set; however, for the patient, flare appears to be a multi-layered and complex phenomenon, set against a background of "normal" fluctuating disease activity. Patients' experiences of their disease in flare may include more than physical symptoms and could encompass systemic features; "flare" may even be preceded by subtle warning signs. Therefore, our data thus far suggest that patient-focused variables encompassing disease worsening may not be adequately reflected within current core sets that are used in randomized controlled trials or observational studies. It is clear that patients' participation will be a critical element to define worsening, to develop acceptable and understandable language to capture flare, and to

test definitions for MDD and MCID. In establishing the "truth" of any definition to capture flare, its face validity requisitely depends on the incorporation of the patient perspective.

There is also a timely and important opportunity to use newly established questionnaires developed by other groups to explore the symptoms most sensitive in detecting a flare. Such efforts are being led through the Patient-Reported Outcomes Measurement Information System (PROMIS) initiative by the US National Institutes of Health (NIH), with additional work ongoing for harmonization through the World Health Organization (WHO) [13,14].

Flares in clinical research

Data from which many initial guidelines are developed and/or validated come from the analysis of cohorts and longitudinal observational studies of patients with RA, increasing their generalizability. Longitudinal observational studies usually provide data on disease improvement using the same outcome measures. However, some may also contain important information to aid in a better understanding of disease worsening, although there are limited published data on disease worsening thus far. In the absence of a gold standard to define a flare, studies have used physician decision to change or add therapy as a surrogate. In one of these studies, Yaziki et al. retrospectively reviewed charts from RA patients and examined the numbers of patients in whom a physician had noted a "flare" or in whom a change in medications was instituted in the context of comparing patients who were receiving etanercept with those who were not. They demonstrated that this drug led to fewer patients having a flare over the period of observation [15]. In another abstract, the CORRONA (Consortium of Rheumatology Researchers of North America, Inc.) database was examined for worsening in core set measurements leading up to a decision to increase therapy [16]. In this study, the sensitivity and specificity of various measurements of disease activity (e.g. patient global assessment and CDAI) were examined. For example, relative to a change in therapy, patient global assessment of disease activity was among the better measures with the highest sensitivity and specificity [16]. A similar approach is being undertaken at King's College London, London, UK, using a database of RA patients attending routine clinics. It examines changes in disease activity measures when patients/clinicians report a "flare" or "disease worsening" as well as factors that are associated with change in therapy (Choy, personal communication). These studies provide some initial insights into understanding "flare" in clinical practice. However, certain difficulties arise in such analyses as they often capture data only intermittently and may fail to capture persistent worsening over time. Additional analysis is needed to better understand whether thresholds for changes in medications may be dependent on the starting point of a patient, or if their disease duration is a potential factor that weighs into a decision to change treatment. Existing longitudinal cohorts are also limited by their potential lack of capture of variables that may distinctively factor into disease worsening as opposed to disease improvement. For example, evaluation of fatigue, which was not included in the original core set of RA measurements, may not be present in all existing RA registries or cohorts. Since the development of the original RA core set, fatigue has been increasingly recognized by patients and clinicians as an important outcome variable to assess both improvement and worsening [17,18].

Defining flare for clinical trials

A particularly critical area in need of a definition of flare is in the field of clinical trials (**Table 1**). Outcomes and endpoints in clinical trials have traditionally reflected improvement from a baseline (e.g. ACR20/50/70 response, or EULAR good responder) or, more recently, endpoints have been reflective of a low disease activity state (e.g. DAS remission). The validation of the component measurements, MDD, and MCID within these indices (e.g. HAQ, patient global assessment, patient pain, SF-36) has been conducted based on improvement. There are only a few examples in adult RA of clinical trials that incorporate a definition of flare or disease worsening [19].

"In the absence of a gold standard to define a flare, studies have used physician decision to change or add therapy as a surrogate"

Within the confines of a clinical trial, patients who worsen or require the addition of DMARDs may be managed in different ways. They may be withdrawn from the protocol with no further data recorded and their status may be reported as withdrawn due to lack of efficacy; they may be rolled over to an active treatment arm; or they may be permitted short courses of steroids or intra-articular injections and continued in the study with an adverse event reported as "RA flare".

In most circumstances in clinical trials, assessments are made based on a single point in time measurement, thus if the patient is in the midst of a temporary and self-reversible oscillation (either good or bad), the data may not reflect the overall trend of their disease activity. Others have documented this as "regression to the mean", and it is an important factor that should be considered when interpreting the results of clinical trials [20,21]. Similarly to the case for improvement, considerations of expected fluctuations will be important to understand in the setting of disease worsening.

The ability to induce significant improvements in disease and "remission" leads to the possibility of tapering or discontinuing

medication while maintaining a low disease activity state. However, the ability to test different withdrawal strategies is without a clear regulatory framework. Induction/withdrawal and induction/maintenance strategies are commonplace in oncological conditions. The opportunity to study these strategies will depend on the development of an outcome definition for flare that reflects an appropriate degree of worsening to differentiate between different treatment arms.

In both juvenile chronic arthritis and in lupus, there are examples of "flare" of disease as an outcome [19]. In juvenile rheumatoid arthritis (JRA) clinical trials, all patients are initially treated with an active medication; subsequently, in only patients who respond, there is a randomized withdrawal in which one group of patients receives continuation of therapy while the other group does not. Survival curves are then used to demonstrate differences in time to flare, and numbers of flares, in both groups [22–24]. In lupus, different flare definitions have also been incorporated into clinical trials. Despite "negative" studies of biological agents in SLE [25,26], a recent early success [27] indicates the importance of developing a reliable definition of flare that can detect change over time. One can envision studies for RA in which patients are induced to a state of low disease activity or remission, and then randomized to different regimens for maintenance of response. Such studies are needed to assist practitioners in the day-to-day management of patients.

"Definition of "flare" as an outcome measure must reflect an appropriate degree of worsening to differentiate between treatment arms"

Studies have been conducted in RA in which there was medication withdrawal and evaluation of time to flare. We have recently summarized the measurements included in a flare definition from these clinical trials [19]. In these studies, there was no consistent definition of flare in terms of component or composite measurements. In several studies, an inversion of improvement criteria has been used to define disease worsening that leads to a change in therapy. In two studies of infliximab, a flare was defined as a ≥50% diminution in a prior improvement in combined swollen and tender joints, which then guided an increase in infliximab dose [28,29]. Other studies have used an inverse of EULAR improvement criteria for response based on DAS to define worsening that would lead to a change in therapy. In one study, the concept of persistence of disease was recognized: patients were required to have a worsening of activity that was observed over more than one visit [30]. A recent consensus statement on the use of rituximab also used similar inverted EULAR criteria to define a flare to guide reinfusion of the drug [9]. While inversion of improvement may seem a reasonable definition of flare, its performance is best when the instrument changes linearly. It is most problematic with percentage changes, for example a 20% decrease in tender joint count from 16 is a reduction by 3.2 while a 20% increase from a baseline of a joint count of 4 is 0.8.

As discussed above, there are significant questions and scenarios in which understanding disease flare could help to guide clinical decision-making (e.g. timing of reinfusion or reinjection of a periodically administered treatment, dose escalation, and comparison of maintenance regimens). The development of a flare definition for clinical trials will open significant opportunities to study such questions of importance, to guide clinical practice and improve patient care.

The analysis of biomarkers and their validation with clinical endpoints is an area of active investigation [31–33]. The traditional framework for biomarker analysis evaluates the baseline level of a marker or an early change from baseline to predict a later clinical response [34]. In the context of disease worsening, there is also an opportunity to examine whether early changes in biochemical markers may precede the development of overt clinical symptoms, thus providing predictive value that may be used to anticipate and to potentially intervene in a manner that would avoid or dampen a flare.

A final area in which a "flare" definition is needed is in the context of adverse event reporting in clinical trials. The OMERACT Drug Safety Group recognized that a standardized flare definition was lacking – but necessary, in order to more accurately capture loss of efficacy withdrawals in clinical trials [35].

Statistical analyses

In accordance with OMERACT methodology, we will develop an outcome measurement of flare for use in clinical trials that will fulfill the necessary criteria of truth, discrimination, and feasibility. As part of this method, the variable(s) that show the most pronounced discriminant capacity will be assessed further via quantification of consistency across studies. In order to do this, our OMERACT Flare Working Group will use receiver operating characteristics (ROC) plots to determine a potential threshold to define flare. Changes in disease activity measures at "flare" or "worsening" visit, and the previous clinic/study visit, will be determined. We will apply an observation-based approach involving analysis of existing RA data for which the clinician has considered or categorized the patient as either having a flare or not, from each of the available proxy variables. The ability to detect and discriminate a "flare effect" in study outcomes will be evaluated using the standardized response means (SRMs: the ratio of the group mean difference to the pooled standard deviation of the mean change scores) [36]. Each available outcome variable will be handled as the SRM with standard

error, for each different RA cohort, then handled using a standard meta-analysis approach [37]. The variable having the lowest I^2 value across studies will be considered the most consistent [38]; the variable with the largest SRM (on average) will be considered the most discriminant [39]. Using simple 2×2 tables with two groups according to the presence or absence of a particular sign or symptoms, we will calculate the proportions of patients who are correctly "diagnosed" by the test. We will calculate various sensitivities and specificities, and investigate the best quantitative threshold defining a flare. In order to handle multiplicity, a ROC curve will be generated based on the full dataset for each of the available databases (i.e. studies) for promising outcome variables that consistently discriminate a flare.

The OMERACT Flare Working Group

The sections above introduced some of the many questions that arise in assessing disease worsening and flare in RA. Answering these questions is not easy, and the answers remain mostly unknown. However, they do demonstrate the importance of additional research in this area. An OMERACT group of interested individuals was established in 2006 to begin examining the area of flare. The initial work of the group has been reported [19]. The group has now grown to more than 30 members who participate in various working groups, regular teleconferences, and face-to-face meetings. An extensive research agenda is underway to address the many aspects of flare with the participation of clinical researchers in rheumatology, epidemiologists, patient representatives, regulators, and pharmaceutical companies, taking advantage of the unique perspectives that each of the participants represents. In the last year, significant progress has been made to evaluate longitudinal observational studies, with the examination of pharmaceutical clinical trial data underway. Multinational patient focus groups are also being conducted to incorporate the critical patient perspective, which may not currently be captured in existing clinical trials or observational studies. It is anticipated that preliminary definitions for flare will be tested against existing datasets in the coming year. The incorporation of new domains and instruments will then need to be tested prospectively in further longitudinal observational studies and clinical trials. Using a data-driven Delphi process, the group, including clinicians and patients, will develop definitions of flare that can be used in future in both clinical trials and for clinical care.

Conclusions

The members of the OMERACT group strongly believe that the time has indeed come for flares of RA to be evaluated in more detail. Only through such an exercise to provide an evidence-based and consensus pilot definition, can we then validate such a measure in multiple contexts (such as in existing longitudinal observational studies and clinical trial databases), then incorporate several such definitions, potentially with new domains or instruments, to capture prospective data from longitudinal observational studies and randomized controlled trials.

Alongside these research goals, there is clearly a need to examine the performance (and potential modification) of such a definition for clinical practice. The ACR/EULAR group is conducting a similar activity to define remission of disease [4]. These different exercises will be complementary and allow us to move forward in our ability to study different methods of drug tapering (e.g. drug A vs. B, dose decrement, frequency decrease), study head-to-head interventions in their ability to induce (and maintain) remission, and to establish recommendations that assist in patient management. Until this work is conducted and studied, we will remain in a largely empirical arena, without agreement on study designs or definitions. The international OMERACT group, currently with representatives from North America, the European Union, Australasia, and, importantly, including from its inception the active participation of patients, is uniquely positioned to conduct this work. Through several groups and subgroups, a series of analyses are underway, with anticipated publications to follow as we move through the process of consensus.

We believe that the time to act is *now* to develop a definition of disease worsening or flare that meets the OMERACT standards of Truth, Discrimination, and Feasibility, recognizing that there may well be modifications of any definition for different purposes. The fact that we have reached a place in the care of patients with RA in which we have a pressing need for a flare definition, we believe is a very good thing. No longer must we only concentrate on improvement from a prior state, but now we raise the bar further toward a goal of therapy to prevent or lessen periods of disease worsening, and we move closer toward permanent remission or even cure of disease as an outcome.

Disclosures

The authors declare no financial disclosures relevant to this manuscript.

References

1. Westhovens R, Robles M, Ximenes AC et al. Clinical efficacy and safety of abatacept in methotrexate-naive patients with early rheumatoid arthritis and poor prognostic factors. *Ann Rheum Dis* 2009; [Epub ahead of print].

2. Wells GA, Boers M, Li T et al. Investigating the validity of the minimal disease activity state for patients with rheumatoid arthritis treated with abatacept. *J Rheumatol* 2009;**36**:260–5.
3. Wells GA, Boers M, Shea B et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;**32**:2016–24.
4. van Tuyl LH, Vlad SC, Felson DT et al. Defining remission in rheumatoid arthritis: results of an initial American College of Rheumatology/European League Against Rheumatism consensus conference. *Arthritis Rheum* 2009;**61**:704–10.
5. Sokka T, Makinen H, Puolakka K et al. Remission as the treatment goal – the FIN-RACo trial. *Clin Exp Rheumatol* 2006;**24**(6 Suppl 43):S74–6.
6. Allaart CF, Goekoop-Ruiterman YP, de Vries-Bouwstra JK et al. Aiming at low disease activity in rheumatoid arthritis with initial combination therapy or initial monotherapy strategies: the BeSt study. *Clin Exp Rheumatol* 2006;**24**(6 Suppl 43):S77–82.
7. Schoels M, Aletaha D, Smolen JS et al. Follow-up standards and treatment targets in rheumatoid arthritis (RA): results of a questionnaire at the EULAR 2008. *Ann Rheum Dis* 2009; [Epub ahead of print].
8. Sesin CA, Bingham CO. Remission in rheumatoid arthritis: wishful thinking or clinical reality? *Semin Arthritis Rheum* 2005;**35**:185–96.
9. Smolen JS, Keystone EC, Emery P et al. Consensus statement on the use of rituximab in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;**66**:143–50.
10. Aletaha D, Strand V, Smolen JS et al. Treatment-related improvement in physical function varies with duration of rheumatoid arthritis: a pooled analysis of clinical trial results. *Ann Rheum Dis* 2008;**67**:238–43.
11. Strand V, Singh J. Newer biologic agents improve health-related quality of life and productivity in rheumatoid arthritis. *Drugs* 2009; In press.
12. Hewlett S, Cockshott Z, Byron M et al. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. *Arthritis Rheum* 2005;**53**:697–702.
13. Stucki G, Cieza A, Geyh S et al. ICF Core Sets for rheumatoid arthritis. *J Rehabil Med* 2004;(44 Suppl):S87– 93.
14. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clin Exp Rheumatol* 2005;**23**(5 Suppl. 39):S53–7.
15. Yazici Y, Erkan D, Kulman I et al. Decreased flares of rheumatoid arthritis during the first year of etanercept treatment: further evidence of clinical effectiveness in the "real world". *Ann Rheum Dis* 2002;**61**:638–40.
16. Furst D, Chang H, Ranganath V et al. Defining change in disease activity parameters associated with loss of response over time in RA patients. *Arthritis Rheum* 2007;**56**(9 Suppl.):S709 Abstr. 1817.
17. Kirwan JR, Minnock P, Adebajo A et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. *J Rheumatol* 2007;**34**:1174–7.
18. Aletaha D, Landewe R, Karonitsch T et al.; EULAR; ACR. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Arthritis Rheum* 2008;**59**:1371–7.
19. Bingham CO 3rd, Pohl C, Woodworth TE et al. Developing a standardized definition for disease "flare" in rheumatoid arthritis (OMERACT 9 Special Interest Group). *J Rheumatol* 2009;**36**:2335–41.
20. Greenwood MC, Rathi J, Hakim AJ et al. Regression to the mean using the disease activity score in eligibility and response criteria for prescribing TNF-alpha inhibitors in adults with rheumatoid arthritis. *Rheumatology (Oxford)* 2007;**46**:1165–7.
21. van Vollenhoven RF, Brannemark S, Klareskog L. Dose escalation of infliximab in clinical practice: improvements seen may be explained by a regression-like effect. *Ann Rheum Dis* 2004;**63**:426–30.
22. Lovell DJ, Giannini EH, Reiff A et al. Etanercept in children with polyarticular juvenile rheumatoid arthritis. Pediatric Rheumatology Collaborative Study Group. *N Engl J Med* 2000;**342**:763–9.
23. Ruperto N, Lovell DJ, Quartier P et al. Abatacept in children with juvenile idiopathic arthritis: a randomised, double-blind, placebo-controlled withdrawal trial. *Lancet* 2008;**372**:383–91.
24. Ruperto N, Lovell D, Goodman S et al. Long-term efficacy and safety of adalimumab in children with juvenile rheumatoid arthritis (JRA): data over two years of treatment in a Phase III study. *Ann Rheum Dis* 2007;**66**(Supp II):Abstr. THU 0195.
25. Merrill J, Burgos-Vargas R, Westhovens R et al. The efficacy and safety of abatacept in SLE: results of a 12-month exploratory study [abstract]. *Arthritis Rheum* 2008;ACR Annual Meeting 2008, San Francisco, CA: Abstract L15.
26. Merrill JT, Neuwelt CM, Wallace DJ et al. Efficacy and safety of rituximab in patients with moderately to severely active systemic lupus erythematosus (SLE): results from the randomized, double-blind Phase II/III study EXPLORER. *Arthritis Rheum* 2008;ACR Annual Meeting, San Francisco, CA: Abstract L12.
27. Human Genome Sciences and GlaxoSmithKline announce positive Phase 3 study results for Benlysta. http://www.hgsi.com/latest/human-genome-sciences-and-glaxosmithkline-announce-positive-phase-3-study-results-for-benl.html (Accessed August 4, 2009).
28. Rahman MU, Strusberg I, Geusens P et al. Double-blinded infliximab dose escalation in patients with rheumatoid arthritis. *Ann Rheum Dis* 2007;**66**:1233–8.
29. Fleischmann RM, Cohen SB, Moreland LW et al. Methotrexate dosage reduction in patients with rheumatoid arthritis beginning therapy with infliximab: the Infliximab Rheumatoid Arthritis Methotrexate Tapering (iRAMT) trial. *Curr Med Res Opin* 2005;**21**:1181–90.
30. Van den Bemt BJ, den Broeder AA, Snijders GF 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.
31. Lassere MN, Johnson KR, Boers M et al. Definitions and validation criteria for biomarkers and surrogate endpoints: development and testing of a quantitative hierarchical levels of evidence schema. *J Rheumatol* 2007;**34**:607–15.
32. Maksymowych WP, Landewe R, Boers M et al. Development of draft validation criteria for a soluble biomarker to be regarded as a valid biomarker reflecting structural damage endpoints in rheumatoid arthritis and spondyloarthritis clinical trials. *J Rheumatol* 2007;**34**:634–40.
33. Keeling SO, Landewe R, van der Heijde D et al. Testing of the preliminary OMERACT validation criteria for a biomarker to be regarded as reflecting structural damage endpoints in rheumatoid arthritis clinical trials: the example of C-reactive protein. *J Rheumatol* 2007;**34**:623–33.
34. Garnero P, Aronstein WS, Cohen SB et al. Relationships between biochemical markers of bone and cartilage degradation with radiological progression in patients with knee osteoarthritis receiving risedronate: the Knee Osteoarthritis Structural Arthritis randomized clinical trial. *Osteoarthritis Cartilage* 2008;**16**:660–6.
35. Woodworth T, Furst DE, Alten R 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.
36. Wells G, Li T, Maxwell L, Maclean R et al. Responsiveness of patient reported outcomes including fatigue, sleep quality, activity limitation, and quality of life following treatment with abatacept for rheumatoid arthritis. *Ann Rheum Dis* 2008;**67**:260–5.
37. Normand SL. Meta-analysis: formulating, evaluating, combining, and reporting. *Stat Med* 1999;**18**:321–59.
38. Higgins JP, Thompson SG, Deeks JJ et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557–60.
39. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? *Stat Med* 2002;**21**:1559–73.