



Felson, D. T., Smolen, J. S., Wells, G., Zhang, B., van Tuyl, L. H. D., Funovits, J., Aletaha, D., Allaart, C. F., Bathon, J., Bombardieri, S., Brooks, P., Brown, A., Matucci-Cerinic, M., Choi, H., Combe, B., de Wit, M., Dougados, M., Emery, P., Furst, D., ... Boers, M. (2011). American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Care and Research*, 63(3), 573-586. <https://doi.org/10.1002/art.30129>

Publisher's PDF, also known as Version of record

Link to published version (if available):
[10.1002/art.30129](https://doi.org/10.1002/art.30129)

[Link to publication record on the Bristol Research Portal](#)
PDF-document

University of Bristol – Bristol Research Portal

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/brp-terms/>

Arthritis & Rheumatism

An Official Journal of the American College of Rheumatology
www.arthritisrheum.org and wileyonlinelibrary.com

American College of Rheumatology/European League Against Rheumatism Provisional Definition of Remission in Rheumatoid Arthritis for Clinical Trials

David T. Felson,¹ Josef S. Smolen,² George Wells,³ Bin Zhang,⁴ Lilian H. D. van Tuyl,⁵ Julia Funovits,⁶ Daniel Aletaha,⁶ Cornelia F. Allaart,⁷ Joan Bathon,⁸ Stefano Bombardieri,⁹ Peter Brooks,¹⁰ Andrew Brown,¹¹ Marco Matucci-Cerinic,¹² Hyon Choi,⁴ Bernard Combe,¹³ Maarten de Wit,¹⁴ Maxime Dougados,¹⁵ Paul Emery,¹⁶ Daniel Furst,¹⁷ Juan Gomez-Reino,¹⁸ Gillian Hawker,¹⁹ Edward Keystone,²⁰ Dinesh Khanna,¹⁷ John Kirwan,²¹ Tore K. Kvien,²² Robert Landewé,²³ Joachim Listing,²⁴ Kaleb Michaud,²⁵ Emilio Martin-Mola,²⁶ Pamela Montie,²⁷ Theodore Pincus,²⁸ Pamela Richards,²⁹ Jeffrey N. Siegel,³⁰ Lee S. Simon,³¹ Tuulikki Sokka,³² Vibeke Strand,³³ Peter Tugwell,³ Alan Tyndall,³⁴ Desirée van der Heijde,⁷ Suzan Verstappen,³⁵ Barbara White,³⁶ Frederick Wolfe,³⁷ Angela Zink,²⁴ and Maarten Boers⁵

This criteria set has been approved by the American College of Rheumatology (ACR) Board of Directors and the European League Against Rheumatism (EULAR) Executive Committee as Provisional. This signifies that the criteria set has been quantitatively validated using patient data, but it has not undergone validation based on an external data set. All ACR/EULAR-approved criteria sets are expected to undergo intermittent updates.

The American College of Rheumatology is an independent, professional, medical and scientific society which does not guarantee, warrant, or endorse any commercial product or service.

Objective. Remission in rheumatoid arthritis (RA) is an increasingly attainable goal, but there is no widely used definition of remission that is stringent but achievable and could be applied uniformly as an outcome measure in clinical trials. This work was undertaken to develop such a definition.

Methods. A committee consisting of members of

the American College of Rheumatology, the European League Against Rheumatism, and the Outcome Measures in Rheumatology Initiative met to guide the process and review prespecified analyses from RA clinical trials. The committee requested a stringent definition (little, if any, active disease) and decided to use core set measures including, as a minimum, joint counts and levels of an acute-phase reactant to define remission. Members were surveyed to select the level of each core set measure that would be consistent with remission. Candidate definitions of remission were tested, including those that constituted a number of individual measures of remission (Boolean approach) as well as definitions using disease activity indexes. To select a definition of remission, trial data were analyzed to examine the added contribution of patient-reported outcomes and the ability of candidate measures to predict later good radiographic and functional outcomes.

This article is published simultaneously in the March 2011 issue of *Annals of the Rheumatic Diseases*.

The views presented in this article do not necessarily reflect those of the United States Food and Drug Administration.

Supported by the American College of Rheumatology, the European League Against Rheumatism, and the NIH (grant AR-47785).

¹David T. Felson, MD, MPH: Boston University School of Medicine, Boston, Massachusetts, and University of Manchester, Manchester, UK; ²Josef S. Smolen, MD: Medical University of Vienna and Hietzing Hospital, Vienna, Austria; ³George Wells, PhD, MSc, Peter Tugwell, MD, MSc, FRCPC: University of Ottawa, Ottawa, Ontario, Canada; ⁴Bin Zhang, ScD: Hyon Choi, MD, DrPH: Boston University School of Medicine, Boston, Massachusetts; ⁵Lilian H. D. van Tuyl, PhD, MSc, Maarten Boers, MD, MSc, PhD: VU University

Results. Survey results for the definition of remission suggested indexes at published thresholds and a count of core set measures, with each measure scored as 1 or less (e.g., tender and swollen joint counts, C-reactive protein [CRP] level, and global assessments on a 0–10 scale). Analyses suggested the need to include a patient-reported measure. Examination of 2-year followup data suggested that many candidate definitions performed comparably in terms of predicting later good radiographic and

functional outcomes, although 28-joint Disease Activity Score–based measures of remission did not predict good radiographic outcomes as well as the other candidate definitions did. Given these and other considerations, we propose that a patient’s RA can be defined as being in remission based on one of two definitions: (a) when scores on the tender joint count, swollen joint count, CRP (in mg/dl), and patient global assessment (0–10 scale) are all ≤ 1 , or (b) when the score on the Simplified Disease Activity Index is ≤ 3.3 .

Conclusion. We propose two new definitions of

Medical Center, Amsterdam, The Netherlands; ⁶Julia Funovits, Dipl. Ing, Daniel Aletaha, MD, MSc: Medical University of Vienna, Vienna, Austria; ⁷Cornelia F. Allaart, MD, PhD, Desirée van der Heijde, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; ⁸Joan Bathon, MD: Johns Hopkins University School of Medicine, Baltimore, Maryland (current address: Columbia University, New York, New York); ⁹Stefano Bombardieri, MD: University of Pisa, Pisa, Italy; ¹⁰Peter Brooks, MD, FRACP: University of Melbourne, Melbourne, Victoria, Australia; ¹¹Andrew Brown, MBChB, MRCP, PhD: University of York, York, UK; ¹²Marco Matucci-Cerinic, MD, PhD: University of Florence, Florence, Italy; ¹³Bernard Combe, MD, PhD: Montpellier University Hospital, Montpellier, France; ¹⁴Maarten de Wit, MSc: Amsterdam, The Netherlands; ¹⁵Maxime Dougados, MD: Paris-Descartes University, UPRES-EA, Assistance Publique Hôpitaux de Paris, and Cochin Hospital, Paris, France; ¹⁶Paul Emery, MA, MD, FRCP: University of Leeds and NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds, UK; ¹⁷Daniel Furst, MD, Dinesh Khanna, MD, MSc: David Geffen School of Medicine at the University of California in Los Angeles; ¹⁸Juan Gomez-Reino, MD, PhD: Hospital Clinico Universitario and Universidad de Santiago de Compostela, Santiago, Spain; ¹⁹Gillian Hawker, MD, MSc, FRCPC: Women’s College Hospital and University of Toronto, Toronto, Ontario, Canada; ²⁰Edward Keystone, MD, FRCPC: University of Toronto, Toronto, Ontario, Canada; ²¹John Kirwan, MD: University of Bristol, Bristol, UK; ²²Tore K. Kvien, MD, PhD: Diakonhjemmet Hospital, Oslo, Norway; ²³Robert Landewé, MD: University Hospital Maastricht, Maastricht, The Netherlands; ²⁴Joachim Listing, PhD, Angela Zink, PhD: German Rheumatology Research Center, Berlin, Germany; ²⁵Kaleb Michaud, PhD: National Data Bank for Rheumatic Diseases, Wichita, Kansas and University of Nebraska Medical Center, Omaha; ²⁶Emilio Martin-Mola, PhD: Hospital Universitario La Paz, Madrid, Spain; ²⁷Pamela Montie: Vancouver, British Columbia, Canada; ²⁸Theodore Pincus, MD: New York University School of Medicine, New York, New York; ²⁹Pamela Richards: Bristol, UK; ³⁰Jeffrey N. Siegel, MD: US Food and Drug Administration, Washington, DC (current address: Genentech, South San Francisco, California); ³¹Lee S. Simon, MD: SDG LLC, Cambridge, Massachusetts; ³²Tuulikki Sokka, MD, PhD: Jyväskylä Central Hospital, Jyväskylä, Finland; ³³Vibeke Strand, MD: Stanford University School of Medicine, Palo Alto, California; ³⁴Alan Tyndall, MD, FRACP: University of Basel, Basel, Switzerland; ³⁵Suzan Verstappen, PhD: University of Manchester, Manchester, UK; ³⁶Barbara White, MD: MedImmune, Gaithersburg, Maryland; ³⁷Frederick Wolfe, MD: National Data Bank for Rheumatic Diseases and University of Kansas, Wichita, Kansas.

Drs. Felson and Smolen contributed equally to this work.

Dr. Smolen has received consulting fees, speaking fees, and/or honoraria from Amgen, Abbott, Centocor, Schering-Plough, Wyeth, Bristol-Myers Squibb, Roche, UCB, and AstaZeneca (less than \$10,000 each); he has received grants from Schering-Plough, Roche, UCB, Bristol-Myers Squibb, and Abbott. Dr. Aletaha has received consulting fees, speaking fees, and/or honoraria from Abbott, Roche, UCB, Bristol-Myers Squibb, Schering-Plough, and Wyeth (less than \$10,000 each). Dr. Allaart has received consulting fees, speaking fees, and/or honoraria from Schering-Plough, Centocor, and UCB (less than \$10,000 each). Dr.

Bathon has received consulting fees, speaking fees, and/or honoraria from Crescendo Biosciences and Roche (less than \$10,000 each); she has received research contracts from Biogen Idec. Dr. Brown has received speaking fees and/or honoraria from Schering-Plough, Abbott, and Pfizer (less than \$10,000 each). Dr. Matucci-Cerinic has provided paid consultation to Pfizer, Actelion, and Schering (less than \$10,000 each). Dr. Choi has received honoraria for serving on advisory boards for TAP Pharmaceuticals and Savient (less than \$10,000 each). Dr. Furst has received consulting fees, speaking fees, and/or honoraria from Abbott, Actelion, Amgen, Bristol-Myers Squibb, Biogen Idec, Centocor, Genentech, Gilead, GlaxoSmithKline, Merck, Nitec, Novartis, UCB, Wyeth, and Xoma (less than \$10,000 each); he has received research funding from Abbott, Actelion, Amgen, Bristol-Myers Squibb, Genentech, Gilead, GlaxoSmithKline, Nitec, Novartis, Roche, UCB, Wyeth, and Xoma. Dr. Gomez-Reino has received consulting fees, speaking fees, and/or honoraria from Schering-Plough, Bristol-Myers Squibb, Wyeth, Roche, and UCB (less than \$10,000 each). Dr. Khanna has received consulting fees, speaking fees, and/or honoraria from Abbott and UCB (less than \$10,000 each). Dr. Landewé has received consulting fees, speaking fees, and/or honoraria from Abbott, Centocor, Schering-Plough, Wyeth, Pfizer, UCB, Merck, Bristol-Myers Squibb, and Amgen (less than \$10,000 each). Dr. Martin-Mola has received consulting fees from Merck, Sharp, and Dohme, Pfizer, and Roche (less than \$10,000 each). Dr. Pincus has received consulting fees, speaking fees, and/or honoraria from Amgen, Abbott, Bristol-Myers Squibb, Centocor, UCB, Wyeth, and Genentech (less than \$10,000 each) and research grants from Amgen, Bristol-Myers Squibb, UCB, and Centocor. Dr. Simon has received consulting fees from Affinergy, AstraZeneca, Abraxis, Alpha Rx, Nuvo/Dimethaid Research, Roche, Pfizer, Novartis, PLx Pharma, Hisamitsu, Dr Reddys, Avanir, Cerimon, Alimera, Paraaxel, Nitec, Bayer, Rigel, Chelsea, Regeneron, Cypress Biosciences, Nicox, Biocryst, Extera, Wyeth, Solace, Puretech-ventures, White Mountain Pharma, Abbott, Omeros, Jazz, Takeda, Teva, Zydus, Proprius, Alder, Cephalon, Seprecor, Purdue, EMD Merck Serono, Altea, Talagen, TiGenix, Antigenics, Forest, Genzyme, CaloSyn, King, Pozen, IL Pharma, Analgesic Solutions, and US WorldMeds (less than \$10,000 each) and from Savient and Horizon (more than \$10,000 each); owns stock or stock options in Savient; and has provided paid consultation to Leerink Swann, Luxor, Nomura, and Fidelity, investment analysis firms. Dr. Sokka has received consulting fees, speaking fees, and/or honoraria from Abbott, Pfizer, and UCB (less than \$10,000 each). Dr. Tugwell has received consulting fees from Bristol-Myers Squibb, Chelsea, and UCB (more than \$10,000 each); he has received research grants from Aventis (HMR), Biomatrix, Cigna, Genzyme, Merck, Novartis, Parke Davis, Pfizer, Rhone-Poulenc, Sandoz, and SmithKline Beecham. Dr. Zink has received speaking fees from Abbott, Bristol-Myers Squibb, Roche, Wyeth, and UCB (less than \$10,000 each).

Address correspondence to Maarten Boers, MD, MSc, PhD, Department of Epidemiology and Biostatistics, VU University Medical Center, PK 6Z 165, PO Box 7057, 1007 MB Amsterdam, The Netherlands. E-mail: eb@vumc.nl.

Submitted for publication June 3, 2010; accepted in revised form October 26, 2010.

remission, both of which can be uniformly applied and widely used in RA clinical trials. We recommend that one of these be selected as an outcome measure in each trial and that the results on both be reported for each trial.

With the advent of new therapies and treatment strategies for rheumatoid arthritis (RA), remission has become a realistic goal (1–3) and has recently become a secondary or even primary end point for clinical studies and trials (4–8). Remission is also regarded as a major therapeutic target in clinical practice (9–12) and can be achieved in a significant proportion of patients receiving routine followup care (13–15). However, the formal definition of RA remission differs between studies.

The current American College of Rheumatology (ACR) definition of remission in RA (16) was developed in 1981, prior to the introduction of the RA core set measures (17). In this classic article, Pinals et al stated: “. . . ‘complete remission’ implies the total absence of all articular and extraarticular inflammation and immunologic activity related to rheumatoid arthritis (RA).” Recognizing that detecting such a state could entail documentation by “extraordinary measures,” they settled on the concept of “complete clinical remission,” aiming to achieve “uniformity in clinical application using generally acceptable and convenient measures.” Even though this concept is of considerable value as a therapeutic target in trials and clinical practice, the 1981 ACR definition has not been widely used in clinical trials in RA because it contains some elements not in the core set (morning stiffness, swelling in tendon sheaths) and a time requirement. Also, this original version was so stringent that few patients met the criteria. Subsequently, many modifications of the ACR criteria were developed, usually omitting one or more of the measures as well as the time requirement.

The development of composite indices of disease activity allowed definition of cut point values representing remission (18–20), but their validation was often limited to comparisons with such modified ACR criteria. For instance, we now know that the widely used definition of remission based on a 28-joint Disease Activity Score (DAS28) (21) of <2.6 (18) better represents minimal disease activity than remission, since multiple joints can remain swollen or tender at that score (19,22–24). This is further exemplified by the fact that in many recent clinical trials the proportion of patients with an ACR70 response (i.e., improvement based on the ACR preliminary definition of improvement [25] but applying 70% improvement instead of 20%) is similar to, or even lower than, the proportion of patients attaining remis-

sion as assessed by the DAS28 (5,26–28). Thus, the 1981 statement by Pinals et al (16) remains relevant today: “Substantial variation appears to exist in the concept of remission within the group of participating rheumatologists.”

In the meantime, effective treatments for RA have led to more exacting criteria for improvement (e.g., ACR 50%, 70%, and even 90% improvement) and have led to recently proposed definitions of minimal disease activity (29). In light of the heterogeneity of definitions of remission, the time has come for consensus on a new, uniform remission definition. Therefore, the ACR and the European League Against Rheumatism (EULAR), together with the Outcome Measures in Rheumatology Initiative (OMERACT), jointly convened a committee to redefine remission in RA. This committee subsequently published a systematic review of the prognostic validity of current remission definitions (30) as well as an outline of the goals of redefining remission and the methods by which the goals would be attained (2).

As noted in this already published outline of our goals (2), the committee decided by consensus to create a stringent definition for remission and agreed that any definition should include, as a minimum, tender and swollen joint counts and levels of an acute-phase reactant. Excluded were treatment, duration of remission (the committee believed this should be specified in each trial report), and measures of physical function and radiographic damage. The latter two were to be used to validate candidate remission definitions: the chosen definition should predict future good functional outcomes and absence of radiographic damage progression. Remission should also predict future remission and minimal disease activity, i.e., show stability. Finally, the requirement for full or 28-joint counts had to be studied. The committee suggested that core set measures should be used to define remission and that any definition of remission in clinical trials should look toward and make possible a similar definition in clinical practice.

The selection of the optimal definition of remission was guided by the research agenda as put forth by the committee at the beginning of our deliberations. In general, the evidence-based consensus method was in accordance with similar activities previously performed by OMERACT as well as ACR and EULAR (31–33) with the intent of deriving a definition that would pass the OMERACT filter of truth, discrimination, and feasibility (34). Herein we present the results of analyses addressing this research agenda, report on later meetings of the committee in which the results were evaluated, and present a consensus definition of remission.

METHODS

General aspects. The initial committee was formed by inviting members of the ACR committee that had previously formulated the new ACR response criteria (35), principal investigators of recent clinical trials, methodologists and patient experts from the OMERACT community, and ACR, EULAR, and OMERACT leaders, with a view to being inclusive and geographically representative. All members present at one of the committee meetings were asked to consider becoming authors of the present report. A patient expert (or research partner) can be described as a patient involved in research based on personal experience of disease that is not available to most researchers, but that complements researchers' analytical skills and scientific perspective (36). During the development of the remission definition, 6 patient experts were involved, 3 of whom (MdW, PM, PR) are authors. Using the specifications provided by the committee, a steering group (DTF, JSS, GW, BZ, LHDvT, JF, MB) designed and performed the necessary investigations; this included a survey as well as analyses of clinical trial data. Clinical trial data banks with slightly different total patient numbers and data composition were created at different centers. Because industry-funded clinical trials have been the largest RA trials carried out and data collected in these trials would be useful for testing hypothesized remission definitions, we solicited industry-funded RA clinical trials data, with the companies' approval. Industry had no role in data analysis, criteria development, testing or evaluating the process, or final choices made by the committee, nor were they consulted or involved in manuscript development. These data were analyzed by the steering group according to the committee's specifications and methodologic discussions in the steering group.

Survey. Our first goal in selecting candidate criteria for remission was to define what cut points in each of the core set measures might constitute remission. In a survey we asked committee members (experienced RA clinical researchers and patient experts) what level of residual activity in individual core set measures would constitute remission. For all measures except joint counts, we used a 0–10 scale. For C-reactive protein (CRP), we used a scale in milligrams per deciliter. For initial analyses of joint counts, we used a 28-joint count and then examined its validity for remission (see below). We asked committee members to state the highest level of each core set measure that would be compatible with remission if it were the only measure assessed, and also asked for the highest level of a particular core set measure that would be compatible with remission if all other measures suggested remission.

Value of patient-reported outcomes. The committee raised the question as to whether patient-reported outcomes should be included in the definition of remission. We addressed this issue by asking whether patient-reported outcomes at the level of remission discriminated between active versus control treatment in trials. In a subset of our data bank, which comprised core set data from 4 clinical trials (37–40), we performed two sets of analyses in each trial. In both analyses the dependent variable was treatment assignment. First, we carried out a logistic regression with each of the core set measures as predictors (recoded as remission level, e.g., swollen joint count ≤ 1 : yes or no). Second, we performed recursive partitioning by classification and regression tree (CART)

analysis on data from the 4 clinical trials, in which we ranked core set measures at remission level based on the tree created from a series of binary splits. Recursive partitioning is a statistical method for multivariable analysis, creating a tree with branches that strives to correctly classify members of the population based on a dichotomous dependent variable. If the patient-reported outcomes helped differentiate active treatment from control (either by being a significant predictor in the regression analysis or by having a high rank in the classification tree), then these outcomes would be said to contribute importantly to defining remission. Patient-reported outcomes tested in this analysis were patient global assessment and patient pain. Functional status measurement was not included, for reasons outlined above.

Assessment of predictive validity. Once we had decided that patient-reported outcomes were to be included and had determined the cut points to be used to define remission, we undertook the analysis of predictive validity. To this end we evaluated various 2-year data sets from randomized clinical trials (patient-level data on 80–90% of patients selected randomly) kindly provided by the sponsors of these studies (35,37,40–43) and obtained permission to use these data for the present analysis. The data are described in more detail in the original publications. For the present analyses, we evaluated only patients for whom all pertinent data over a 2-year period were available.

We initially defined good outcome for radiographic damage and physical function separately. For radiographic damage, the definition comprised stable radiography scores over 1 year (defined as change of ≤ 0 in Sharp scores [44] or modified Sharp/van der Heijde scores [45] during the second year of the trial). For physical function, it comprised stable and low scores on the Health Assessment Questionnaire (HAQ) (46) (change of ≤ 0 and HAQ score consistently ≤ 0.5 during the second year of the trial). We then tested whether patients who met a particular definition of remission at 6 months or 12 months were more likely to have a good outcome in the subsequent period, i.e., between 1 and 2 years after trial initiation. Likelihood ratios were used to compare the proportion of patients having the good outcome whose RA was in remission to the proportion of patients having the good outcome whose RA was *not* in remission. To rank candidate definitions of remission, we used the *P* value from the logistic regression chi-square test. As has been reported (47), most patients in trials who are followed up long term do not show radiographic progression. This limited our capacity to discriminate between candidate definitions of remission. Moreover, intensive therapy with tumor necrosis factor (TNF) inhibitors plus methotrexate (MTX) dissociates clinical disease activity from progression of joint damage, since—unlike patients treated with MTX alone—those receiving aggressive treatments have no or minimal radiographic progression irrespective of their disease activity (48–50). Therefore, we primarily performed the analyses on patients treated with MTX monotherapy; however, we also evaluated TNF inhibitor monotherapy and combination therapy in sensitivity analyses. To assess the robustness of the results, we also performed the analyses on a subset of trial patients with an especially poor prognosis in terms of radiographic disease—i.e., presence of rheumatoid factor *and* presence of radiographic damage at baseline. Finally, we tested an additional definition of a good

outcome, i.e., stability of both radiographic damage and HAQ score.

Selection of candidate definitions. Candidate definitions of remission were selected from two general categories: first, indices that have been widely used, including the DAS28, the Simplified Disease Activity Index (SDAI), and the Clinical Disease Activity Index (CDAI) (18,19,21,32,33,51,52), and second, definitions including one or more core set measures at cut points previously defined by the survey, but requiring all included measures to be at or below that cut point. For example, to meet remission defined as low scores on tender and swollen joint counts and physician and patient global assessments, the patient must have had low scores on all 4 measures. These measures are referred to below as the Boolean measures based on their approach, which is to define each core set measure as in remission or not (values of 0 and 1) and use possible combinations of the patient's core set measure remission status to determine the patient's overall remission status (also 0 or 1).*

Further evaluations of candidate definitions including assessment of face validity. After completing the analysis of predictive validity, we tested our candidate definitions for face validity. Since we had decided that any definition of remission must be stringent with respect to not allowing much disease activity, we studied whether patients could meet a definition of remission yet still have moderate to high levels of disease activity in any core set measure. To do this, in the group of patients meeting a certain definition of remission, we studied the 90th percentile and maximum level of disease activity observed in each core set measure. Last, we looked at recent

trial data to determine what proportion of patients met each remission definition. It was our goal not to have an undetectably low percentage of patients meeting the definition of remission, or one so high as to be unreasonable given clinical experience with these treatments.

We also examined two related issues for our candidate definitions. First, we wished to select a definition(s) that was reliable, and we determined this by analyzing, in one trial with monthly visits, whether a patient whose RA was defined as being in remission at one visit attained the same status at adjacent visits ≤ 1 month from the first; if the disease was not in remission at the adjacent visit, we assessed whether disease activity remained at minimal levels (29). Second, we were concerned that a 28-joint count might not capture actively involved joints outside these 28; to address this, we reviewed literature and analyzed trial data to determine whether we should define remission differently when using 28 versus, for example, 66 joints. For the latter, we evaluated data from a set of trials that included tenderness and swelling counts of individual joints. In these we assessed residual disease activity in ankles and feet in patients with 28-joint counts of ≤ 1 and determined what proportion of such patients would satisfy the other requirements of our candidate definitions. These patients would represent real misclassification ("false-positive" remissions). In the same data set we subsequently investigated whether such misclassification could materially affect the predictive validity of the remission definitions. For this purpose we compared the prevalence of good outcome (damage or function) in patients with "true remission" (i.e., based on full joint counts at ≤ 1) with that in all patients with remission based on 28-joint counts (i.e., "true" plus "false-positive" remissions).

RESULTS

Survey. Twenty-seven committee members, including two patients, completed the survey on threshold levels for remission (Table 1). In the scenario in which

*Boolean measure is the logic that computers use to determine if a statement is true or false. There are 4 main Boolean operators: AND, NOT, OR, and XOR (exclusive OR). Below is an example, from defining remission, of how one operator works:

Assume that x and y are both core set variables for RA whose values are in the range of remission. x AND y returns True if both x and y are true; otherwise the expression returns False. False means that patient's RA is NOT in remission.

Table 1. Threshold levels for remission in the RA core set measures according to the survey of committee members*

Core set measure	Highest level of the core set measure that would be compatible with remission									
	If it were the only measure assessed					If all other measures suggested remission				
	Mean \pm SD	Minimum	Median	80%	Maximum	Mean \pm SD	Minimum	Median	80%	Maximum
TJC28	1.1 \pm 1.3	0	1	2	6	2.6 \pm 2.0	1	2	4	10
Full TJC (68 joints)	1.6 \pm 1.5	0	2	2	6	2.6 \pm 2.0	1	2	4	10
SJC28	0.5 \pm 0.9	0	0	1	4	1.3 \pm 1.3	0	1	2	6
Full SJC (66 joints)	0.6 \pm 0.9	0	0	1	4	1.4 \pm 1.2	0	1	2	6
ESR, mm/hour	21 \pm 6	10	20	25	30	25 \pm 6	20	25	30	40
CRP, mg/dl	0.9 \pm 0.4	0	1	1	2	1.1 \pm 0.6	0	1	1.5	2
Pain, 0–10 scale	1.3 \pm 0.7	0	1	2	3	2.4 \pm 1.3	1	2	3	6
PhGA, 0–10 scale	1.0 \pm 0.9	0	1	1	4	1.6 \pm 1.0	0	2	2	4
PtGA, 0–10 scale	1.2 \pm 0.8	0	1	2	3	2.2 \pm 1.3	0	2	3	6
HAQ, 0–3 scale	0.7 \pm 0.7	0	0.5	0.5	3	0.9 \pm 0.8	0.2	0.6	1	3

* Twenty-seven committee members responded to the survey (25 experienced rheumatoid arthritis clinical researchers and 2 patients). RA = rheumatoid arthritis; 80% = 80th percentile; TJC28 = tender joint count using 28 joints; SJC = swollen joint count using 28 joints; ESR = erythrocyte sedimentation rate; CRP = C-reactive protein; PhGA = physician/observer global assessment; PtGA = patient global assessment; HAQ = Health Assessment Questionnaire.

Table 2. Predictive validity of candidate remission definitions for good outcome in radiographic damage*

Candidate remission definition	Prevalence of good outcome in patients		Positive likelihood ratio (95% CI)	P†
	In remission	Not in remission		
TJC28, SJC28, CRP ≤1	69 (34/49)	50 (154/306)	2.0 (1.1–3.6)	0.01
TJC28, SJC28, CRP, PhGA ≤1	76 (26/34)	51 (162/320)	2.9 (1.3–6.2)	0.004
TJC28, SJC28, CRP, PtGA ≤1	77 (23/30)	51 (165/325)	2.9 (1.3–6.6)	0.006
TJC28, SJC28, CRP, pain ≤1	74 (23/31)	51 (165/324)	2.6 (1.2–5.6)	0.01
TJC28, SJC28, CRP, PhGA, PtGA ≤1	77 (20/26)	51 (168/328)	2.9 (1.2–7.2)	0.01
TJC28, SJC28, CRP, PhGA, pain ≤1	77 (20/26)	51 (168/328)	2.9 (1.2–7.2)	0.01
TJC28, SJC28, CRP, PtGA, pain ≤1	76 (22/29)	51 (166/326)	2.8 (1.2–6.4)	0.001
TJC28, SJC28, CRP, PhGA, PtGA, pain ≤1	76 (19/25)	51 (169/329)	2.8 (1.1–6.8)	0.02
DAS28 <2.6	60 (21/35)	59 (93/157)	1.0 (0.6–1.9)	0.93
DAS28 <2.0	70 (7/10)	59 (107/182)	1.6 (0.4–6.0)	0.48
SDAI ≤3.3‡	77 (27/35)	50 (161/319)	3.0 (1.4–6.4)	0.003
Definitions without CRP (for clinical practice)				
TJC28, SJC28, PhGA, PtGA ≤1	75 (24/32)	51 (167/326)	2.6 (1.2–5.7)	0.01
TJC28, SJC28, PtGA ≤1	75 (27/36)	51 (164/323)	2.6 (1.3–5.4)	0.007
CDAI ≤2.8§	75 (27/36)	51 (164/322)	2.6 (1.3–5.4)	0.006

* Values in the first 2 columns are percentages, with absolute proportions shown in parentheses. Presence or absence of remission, defined according to the given candidate definition, was measured at 6 months after baseline, using combined data from methotrexate-alone treatment groups in 3 trials (37,40,41) (limited to patients with complete data over 2 years). Good radiographic outcome was defined as a change of ≤0 in Sharp/van der Heijde scores between 12 and 24 months after baseline. CRP is in mg/dl. 95% CI = 95% confidence interval; DAS28 = 28-joint Disease Activity Score (see Table 1 for other definitions).

† From chi-square analysis using logistic regression, in which the independent variable was remission (based on the given candidate definition) and the dependent variable was radiographic stability.

‡ The Simplified Disease Activity Index (SDAI) is the simple sum of the TJC (using 28 joints), SJC (using 28 joints), patient global assessment (0–10 scale), physician global assessment (0–10 scale), and CRP level (mg/dl).

§ The Clinical Disease Activity Index (CDAI) is the same as the SDAI, except CRP is not included.

only one variable was available, the responses clustered around core set disease activity levels of 1, such that, for example, the swollen or tender joint count should be 1 or less, the CRP level should be 1 mg/dl or less, and patient and physician global assessments as well as patient pain assessment should be 1 or less on a 10-point scale. The question on which was the highest level of a particular core set measure compatible with remission if all other measures suggested remission yielded more varied answers, with thresholds ranging from 2 for swollen joint count (SJC) and CRP level to 4 for tender joint count (TJC). Since this did not provide us with a single threshold value that was uniform across core set measures, we focused on the more stringent cut points.

Patient-reported outcomes. We then proceeded with an analysis of clinical trial data on active treatment versus control to help determine whether patient-reported outcomes, namely patient global assessment or patient-reported pain, should be incorporated into our definition of remission. In an analysis of 4 clinical trials, both logistic regression and CART analysis demonstrated that these measures added important information to physician-linked measures. In other words, in

these trials, patient global assessment and patient-reported pain were statistically significant predictors that discriminated between treatments after controlling for physician-reported measures (TJC and SJC) and a laboratory measure (CRP). For example, in the CART analysis, among the 4 trials, patient global assessment was the best predictor of treatment assignment among all outcomes in one trial and the fourth best of core set measures in another. Patient-reported pain was the second best predictor (SJC was the best) in a third trial.

Based on these preliminary analyses, we developed a list of candidate remission definitions to test for predictive validity. When presented with the more stringent definitions versus the more relaxed definitions, our committee selected those in the more stringent category and as a consequence, we present results only for these. In accordance with the committee's charge and the assessment of the contribution of patient-reported outcomes, we mainly focused on measures that comprised TJC, SJC, CRP level, and patient global assessment. We tested combinations of these and other core set measures to determine if any group of measures would have important advantages.

Table 3. Predictive validity of candidate remission definitions for good outcome in both radiographic damage and HAQ*

Candidate remission definition	Prevalence of good outcome in patients		Positive likelihood ratio (95% CI)	P†
	In remission	Not in remission		
TJC28, SJC28, CRP ≤ 1	46 (22/48)	17 (51/301)	3.2 (1.9–5.3)	<0.0001
TJC28, SJC28, CRP, PhGA ≤ 1	55 (18/33)	17 (55/315)	4.5 (2.4–8.5)	<0.0001
TJC28, SJC28, CRP, PtGA ≤ 1	66 (19/29)	17 (54/320)	7.2 (3.5–14.8)	<0.0001
TJC28, SJC28, CRP, pain ≤ 1	60 (18/30)	17 (55/319)	5.7 (2.9–11.2)	<0.0001
TJC28, SJC28, CRP, PhGA, PtGA ≤ 1	68 (17/25)	17 (56/323)	8.0 (3.6–17.8)	<0.0001
TJC28, SJC28, CRP, PhGA, pain ≤ 1	64 (16/25)	18 (57/323)	6.7 (3.1–14.5)	<0.0001
TJC28, SJC28, CRP, PtGA, pain ≤ 1	64 (18/28)	17 (55/321)	6.8 (3.3–14.1)	<0.0001
TJC28, SJC28, CRP, PhGA, PtGA, pain ≤ 1	67 (16/24)	18 (57/324)	7.5 (3.4–16.9)	<0.0001
DAS28 <2.6	38 (13/34)	18 (28/154)	2.2 (1.2–4.0)	0.01
DAS28 <2.0	56 (5/9)	20 (36/179)	4.5 (1.3–15.9)	0.01
SDAI ≤ 3.3 ‡	56 (19/34)	17 (54/314)	4.8 (2.6–8.9)	<0.0001
Definitions without CRP (for clinical practice)				
TJC28, SJC28, PhGA, PtGA ≤ 1	68 (21/31)	17 (53/321)	7.9 (3.9–16.0)	<0.0001
TJC28, SJC28, PtGA ≤ 1	66 (23/35)	16 (51/318)	7.2 (3.8–13.9)	<0.0001
CDAI ≤ 2.8 §	63 (22/35)	16 (52/317)	6.4 (3.4–12.0)	<0.0001

* Values in the first 2 columns are percentages, with absolute proportions shown in parentheses. Presence or absence of remission, defined according to the given candidate definition, was measured at 6 months after baseline, using combined data from methotrexate-alone treatment groups in 3 trials (37,40,41) (limited to patients with complete data over 2 years). Good radiographic outcome was defined as a change of ≤ 0 in Sharp/van der Heijde scores between 12 and 24 months after baseline; good outcome on the HAQ was defined as a change of ≤ 0 and a score of ≤ 0.5 at both the 12-month and 24-month time points. CRP is in mg/dl. 95% CI = 95% confidence interval; DAS28 = 28-joint Disease Activity Score (see Table 1 for other definitions).

† From chi-square analysis using logistic regression, in which the independent variable was remission (based on the given candidate definition) and the dependent variable was the combination of radiographic and HAQ stability.

‡ The Simplified Disease Activity Index (SDAI) is the simple sum of the TJC (using 28 joints), SJC (using 28 joints), patient global assessment (0–10 scale), physician global assessment (0–10 scale), and CRP level (mg/dl).

§ The Clinical Disease Activity Index (CDAI) is the same as the SDAI, except CRP is not included.

Predictive validity. We then tested whether patients whose RA was in remission according to one of these definitions had a higher likelihood of a good outcome. We focused on patients receiving MTX monotherapy, although we obtained similar results (not shown) when we analyzed data from all patients. We found that patients whose RA was in remission by several of the Boolean candidate definitions, as well as by the traditional SDAI definition (≤ 3.3) and CDAI definition (≤ 2.8), had an increased likelihood of radiographic stability during the subsequent year (Table 2). However, this was not the case for the DAS28 definition, either at the traditional cut point (< 2.6) or at a more stringent cut point (< 2.0). In contrast, being in remission by any of the definitions increased the likelihood of stability on HAQ scores, without important differences between definitions (data not shown). When we defined good outcome as the combination of radiographic and HAQ stability, we again found that being in remission by any of the candidate definitions increased the likelihood of a good outcome (Table 3). As expected, the performance of the DAS28 at either cutoff was not as good as that of the other definitions. Similar data

were also obtained in an additional data set from the COBRA (Combinatietherapie Bij Reumatoïde Artritis) study (42) (data not shown). However, reaching remission according to the DAS28, both at the traditional cut point (< 2.6) and at a more stringent cut point (< 2.0), was associated only with the likelihood of HAQ stability, and not radiographic stability. Candidate definitions of remission did not differ in their prediction of HAQ stability (data not shown). Additional definitions were tested, including incorporating either remission-level pain or patient global assessment and other variations, and results were similar. Apart from the DAS28 result, the analyses did not help to distinguish between definitions. This was also the case in the analysis using a more strict definition of good outcome, and when we studied only patients with a poor prognosis (data not shown).

Face validity. Face validity of the different candidate definitions, expressed as residual disease activity in the presence of remission, is shown in Table 4. For the Boolean definitions, the high values, as expected, tended to be for core set measures that were not prespecified by the rule. For example, if we used the definition of TJC, SJC, CRP, and pain all ≤ 1 , we found that 10% of

Table 4. Face validity expressed as residual disease activity in the presence of remission*

Candidate remission definition	TJC28		SJC28		CRP		PhGA		PtGA		Pain	
	90%	Max	90%	Max	90%	Max	90%	Max	90%	Max	90%	Max
TJC28, SJC28, CRP ≤1	1	1	1	1	0.6	1	2	6	4	8	4	8
TJC28, SJC28, CRP, PhGA ≤1	1	1	1	1	0.6	1	1	1	2	7	2	8
TJC28, SJC28, CRP, PtGA ≤1	1	1	1	1	0.6	1	2	2	1	1	2	3
TJC28, SJC28, CRP, pain ≤1	1	1	1	1	0.6	1	2	4	2	6	1	1
TJC28, SJC28, CRP, PhGA, PtGA ≤1	1	1	1	1	0.7	1	1	1	1	1	1	3
TJC28, SJC28, CRP, PhGA, pain ≤1	1	1	1	1	0.7	1	1	1	1	5	1	1
TJC28, SJC28, CRP, PtGA, pain ≤1	1	1	1	1	0.7	1	1	2	1	1	1	1
TJC28, SJC28, CRP, PhGA, PtGA, pain ≤1	1	1	1	1	0.7	1	1	1	1	1	1	1
DAS28 <2.6	2	7	4	21	0.7	2.5	2	5	3	8	2	10
DAS28 <2.0	0	3	2	6	0.7	2.5	2	3	2	4	2	4
SDAI ≤3.3	1	2	1	2	0.7	2.7	2	2	1	2	1	3

* Values are the upper limits of residual disease activity in the RA core set measures for candidate definitions of remission observed in trial data sets using all trial arms (methotrexate monotherapy, tumor necrosis factor inhibitor monotherapy, and combination therapy with tumor necrosis factor inhibitors plus methotrexate). CRP is in mg/dl. 90% = 90th percentile; Max = maximum observed value; DAS28 = 28-joint Disease Activity Score; SDAI = Simplified Disease Activity Index (see Table 1 for other definitions).

patients (90th percentile) had physician and patient global assessment scores compatible with active disease. If we used TJC, SJC, and CRP all ≤1, then the patient-reported outcomes often suggested high levels of symptoms. For the traditional DAS28 definition (<2.6), we found that many of the core set measures remained at levels that would be incompatible with remission. This was even the case for DAS28 <2.0, which was a threshold few patients reached. It was not the case for other index measures that defined remission, such as the SDAI or CDAI, where results were closely aligned with the Boolean definitions and the results of our survey.

When we examined the proportion of patients in trials who met candidate definitions of remission (Table 5), we felt that the high prevalence of remission according to the current DAS28 definition lacked face validity. Otherwise, 18–26% of patients receiving combination therapy with TNF inhibitors and MTX met most of these

definitions, compared to only 5–10% of those receiving either monotherapy. We believe these percentages reflect face validity.

Consensus activity. The committee met prior to the ACR Annual Scientific Meeting in October 2009 to discuss the analyses described above. As noted, the committee did not select, in any case, a more relaxed definition of remission, consistent with its earlier directive. During the committee meeting two subgroups were formed to discuss the tabular results presented, especially including results regarding predictive validity. Both groups voted that there should be both a Boolean approach and an index-based definition. One group voted among individual definitions of remission, and in doing so, the highest vote was received for the Boolean definition that included TJC, SJC, CRP, and patient global assessment, all at levels ≤1. The index definition with the highest vote count was SDAI ≤3.3. In the other

Table 5. Face validity expressed as the prevalence of remission (%) in recent trials of patients with rheumatoid arthritis*

Candidate remission definition	DMARD monotherapy (n = 380)	Biologic monotherapy (n = 520)	Combination therapy (n = 330)	Total (n = 1,230)
TJC, SJC, CRP, PtGA ≤1	9	7	22	12
TJC, SJC, CRP, PtGA, pain ≤1	8	6	20	12
TJC, SJC, CRP, PhGA, PtGA ≤1	8	7	20	10
TJC, SJC, CRP, PhGA, pain ≤1	8	6	20	10
TJC, SJC, CRP, PhGA, PtGA, pain ≤1	7	6	18	9
DAS28 <2.6	19	17	35	21
DAS28 <2.0	5	8	24	10
SDAI ≤3.3	10	8	26	14

* From pooled data from refs. 37, 40, and 41. CRP is in mg/dl. DAS28 = 28-joint Disease Activity Score; SDAI = Simplified Disease Activity Index (see Table 1 for other definitions).

subgroup, after a discussion involving all study group members, the same conclusion was reached without a formal vote.

Members of this subgroup noted that in the clinical setting an acute-phase response measure is often not available at every visit and the subgroup suggested that a definition of remission be developed for clinic-based practice that would not require an acute-phase reactant, as long as it would capture remission as stringently as the measure used for clinical trials. Indeed, a Boolean measure comprising TJC, SJC, and patient global assessment provided statistical results similar to those obtained with the same measures encompassing CRP and those obtained with the CDAI, which does not include CRP (Table 2). Thus, these definitions of remission may be used in clinical practice until better measures for that purpose become available.

In a trial with monthly visits we found that our selected definitions of remission showed good reliability. Specifically, among patients whose RA was in remission at one time point, the disease remained in remission 1 month later in 66%, and all the rest met criteria for minimal disease activity (29).

Joint counts. We consulted published literature and our own data analysis to determine if remission thresholds for 28-joint counts should be the same as thresholds for counts with more joints assessed (such as 66 or 68 joints). One study (53) examined whether adding ankles and metatarsophalangeal joints to the 28-joint count affected remission and showed that <10% of patients with no tender or swollen joints in a 28-joint count had tender or swollen ankles or metatarsophalangeal joints and that the average patient global assessment score in these latter patients was significantly higher, suggesting that they would not meet proposed definitions of remission. Landewé and colleagues (24) also noted that defining a patient's disease as being in remission using a 28-joint count often concealed active joints elsewhere, especially in the feet and ankles. However, they also reported that global assessments for patients who had 28 joints in remission but actively involved joints elsewhere resembled those for patients whose disease was not in remission based on a 28-joint count, suggesting that requiring a low patient global assessment score will, to some extent, mitigate the limitation of using a 28-joint count.

In the two trials in our data set that included counts of individual joints with tenderness and swelling, remission prevalence using 66 and 68 joints was 4% and 9%, respectively. As in the studies cited above, we found that patients with 28-joint counts ≤ 1 often had residual tenderness or swelling in the ankles or feet. However,

Table 6. American College of Rheumatology/European League Against Rheumatism definitions of remission in rheumatoid arthritis clinical trials*

Boolean-based definition:

At any time point, patient must satisfy all of the following:

Tender joint count ≤ 1 †

Swollen joint count ≤ 1 †

C-reactive protein ≤ 1 mg/dl

Patient global assessment ≤ 1 (on a 0–10 scale)‡

Index-based definition:

At any time point, patient must have a

Simplified Disease Activity Index score of ≤ 3.3 §

* See text and Tables 2 and 3 for recommendations regarding assessment of remission in clinical practice settings.

† For tender and swollen joint counts, use of a 28-joint count may miss actively involved joints, especially in the feet and ankles, and it is preferable to include feet and ankles also when evaluating remission.

‡ For the assessment of remission we suggest the following format and wording for the global assessment questions. Format: a horizontal 10-cm visual analog or Likert scale with the best anchor and lowest score on the left side and the worst anchor and highest score on the right side. Wording of question and anchors: For patient global assessment, “Considering all of the ways your arthritis has affected you, how do you feel your arthritis is today?” (anchors: very well—very poor). For physician/assessor global assessment, “What is your assessment of the patient's current disease activity?” (anchors: none—extremely active).

§ Defined as the simple sum of the tender joint count (using 28 joints), swollen joint count (using 28 joints), patient global assessment (0–10 scale), physician global assessment (0–10 scale), and C-reactive protein level (mg/dl).

most of these patients did not satisfy the other requirements of our candidate definitions of remission. Nevertheless, the estimates of remission prevalence increased to 6% and 14%, respectively, of the total population when 28-joint counts were used. In another data set from two trials with 2-year followup data, we compared patients whose RA was in remission according to full (66/68) joint counts versus those whose RA was in remission according to only 28-joint counts (i.e., with residual disease activity in joints not assessed). Among the patients with “full joint count remission,” 80% had good outcomes in terms of radiographic damage (no change in Sharp score); this number decreased by 1% among patients with only “28-joint count remission.” Likewise, among the patients with “full joint count remission,” 90% had good HAQ outcomes; this number decreased by 1% and 4%, respectively, in the 2 trials, among patients with only “28-joint count remission.” Based on these analyses we concluded that the overall impact (misclassification) due to reduced joint counts is small.

The final recommended definitions of remission are presented in Table 6. Specific suggestions on how

to measure components of the definitions are also provided.

DISCUSSION

Based on considerations of face and predictive validity, the need for stringency, and the need to include patient-reported outcomes, the ACR/EULAR committee charged with defining remission in RA has produced two definitions for evaluating remission in clinical trials. One is a Boolean-based definition, more categorical in structure than the traditional definition from Pinals et al (16), and the other is based on a composite index of RA activity, the SDAI (19,51).

Ideally, we would have liked to select a candidate definition that clearly differentiated patients whose long-term course was without disease progression versus those whose disease continued to progress. Our analysis of long-term data confirmed the findings of our systematic review (30) that most definitions of remission did well—i.e., that patients whose RA was in remission at any point during a clinical trial, based on any of the definitions we used, were likely to have long-term courses that were better than those of patients who did not meet the definition of remission. One exception was the DAS28, which has been shown previously to allow for significant residual disease activity (19,22,23,54,55). Thus, except for DAS28-based definitions, differences in predictive validity between candidate definitions were small (see Tables 2 and 3), and it was difficult to differentiate the course of patients meeting any of these definitions of remission. Among the many definitions tested, none importantly exceeded the ability of the ultimately selected criteria to predict favorable long-term effects on radiographic progression and physical function. Although we can confirm the predictive validity of remission, the goal of the work was to define remission, not to develop a predictive marker.

In our data sets we assessed definitions of remission by 28-joint counts. When we examined more comprehensive counts among patients with disease remission in the 28 joints, we found that residual disease activity was frequently present in ankles and feet. However, most of these patients failed to meet other criteria in the remission definition (e.g., their patient global assessments were often high). In other words, even when joints other than the 28 joints counted were swollen or tender, other measures of disease activity often prevented misclassification of these patients as having disease in remission. In addition, the impact of misclassification on long-term outcome proved to be small. We should also

bear in mind that the assessment of ankles and forefeet is particularly limited and poorly reproducible (56). In line with this, the discordance between tenderness and swelling has proved to be greater in the joints of the feet than in other joints (57). Therefore, we do not *require* inclusion of ankles and forefeet in the assessment of remission but *recommend* that these joints are also included in the examination. Investigators should always report which joints were examined.

In 2008, EULAR and the ACR recommended that in each RA trial, the percentage of patients achieving a low disease activity state and remission should be reported (32,33). On the basis of the present analyses and consensus, we suggest that remission based on one of the definitions recommended here be reported as a preselected outcome measure in trials, and that results for both be included in trial reports. Of the approaches to defining low disease activity, the OMERACT definitions of “minimal disease activity,” designed to reflect the “next best” option apart from remission, have been the best vetted and were consensually developed (29).

There are a few limitations to our approach, and possibly to the definitions produced as a consequence. First, we used a HAQ score of ≤ 0.5 as evidence for stability of the remission criteria; while this is a disability score that is essentially above values obtained in the general population (58), many of the studies evaluated were of patients with longstanding disease who are known to accumulate significant irreversible disability (59). However, we accounted for this potential contrast to the normal situation by also requiring that HAQ scores did not deteriorate at all over a full 1-year period (HAQ change ≤ 0 during the second year of observation).

Second, we have not yet validated the recommended definitions of remission in observational data sets. This is the next step in our work. In developing definitions, we anticipated clinic-based evaluations, trying to choose definitions of remission that would be easy to apply in an observational context and take advantage of variables that are probably already being measured. In clinical practice, data on acute-phase reactants are frequently not immediately available, and therefore, an additional set of a Boolean definition and an index-based definition not requiring acute-phase reactants is provided for that setting. Nevertheless, our preliminary suggestions for defining remission in clinical practice are still incomplete, as we did not test them in a clinic-based setting. While the remission definitions not requiring an acute-phase reactant performed comparably with those that do require this parameter, the committee believes that including an acute-phase reactant for re-

porting remission in clinical trials is preferable because acute-phase reactants are important predictors of later radiographic damage (60–62).

Another limitation of the proposed definition is that the patient experience of the remission may not have been adequately captured with only one element, the patient's global assessment of his or her disease activity. Indeed, an index based on patient measures alone may clinically discriminate between active and control treatment as well as do some of the indexes tested in this effort (63,64). However, the committee had stipulated that joint counts should be part of the remission criteria; moreover, joints are the "organ" involved in RA, and in the context of assessing remission it was deemed advisable to assess that organ.

Further, fatigue was not evaluated (65). However, fatigue was not assessed in most trials published over the last decade, including those used here for the derivation of the remission criteria. We were also unable to procure data sets that contained information on other non-core set measures. As these data sets are likely to become available only over the course of several years, we decided not to postpone the development of the new remission definition. Indeed, we believe it was important to spend more time developing the concept of patient-assessed "absence of disease." This will require qualitative research involving focus groups, as well as quantitative research, e.g., collection of patient-related outcome data in clinical trials, a task that will be taken forward within the OMERACT framework. Once a working definition of this concept is available, it can be compared with the proposed definition of remission.

Yet another theoretical limitation is that imaging results are not included in our definition of remission. Our goal was to use clinical parameters that are widely used and convenient to assess, but we recognize that residual synovitis may exist in many patients whose disease appears inactive based on conventional clinical evaluation (55,66,67). Importantly, however, our definitions of remission were associated with a retardation of radiographic progression, suggesting that the clinical definition has biologic meaning. Moreover, findings of a recent sonographic analysis of RA patients whose disease was in remission as defined by different means (55) were in accordance with the present results. Thus, while our definitions permit a tender or swollen joint to be present, we require multiple pieces of evidence of inactive disease (1 or no tender or swollen joint, low acute-phase reactant level, and assessment by the patient that the disease is inactive) before a patient meets remission criteria. Since inactive disease may be accom-

panied by 1 residual swollen or tender joint and since the reliability of the examination diminishes with the number of joints with active disease, this procedure enhances the sensitivity of our definition of remission.

We should note that the trial data sets we tested included CRP more frequently than erythrocyte sedimentation rate (ESR), explaining why our definitions present thresholds for CRP. A similar ESR threshold for inactive disease might be <20 mm/hour for men and <30 mm/hour for women, or even lower, but this may require further testing (68). Our preference for CRP is in part because it can be standardized across centers, making it the preferred acute-phase reactant measure in multicenter trials. Also, while CRP levels may have different upper limits of normal in different laboratories, the test is widely standardized today, and a value of 1 mg/dl covers all of these upper limits; at 1 mg/dl or less the progression of joint damage is minimized (60,61). Given these findings, the practicality of using the same value, i.e., 1, for all measures was deemed more important than searching for potential minimal differences between cut points of 1 mg/dl or slightly less.

A "treat to target" approach may yield better outcomes than a conventional approach to therapy for RA, and remission can serve as that target for some patients. However, remission according to the stringent definition presented here may not yet be a realistic goal for most patients (10).

In conclusion, we present new definitions of remission for use as outcome measures in RA clinical trials: either the compilation of 4 individual measures or an index-based alternative. We hope that these new definitions will be adopted widely and can provide a uniform approach to assessing this increasingly important outcome.

ACKNOWLEDGMENTS

We are grateful to several colleagues who participated in the committee discussions but declined to be named as coauthors on this report, and to Abbott, Amgen, and Wyeth for sharing clinical trial data.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Boers had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Felson, Smolen, Wells, Zhang, van Tuyl, Allaart, Bathon, Brooks, Brown, Matucci-Cerinic, Choi, Combe, Dougados, Emery, Furst, Gomez-Reino, Hawker, Kirwan, Kvien, Lan-

dewé, Listing, Martin-Mola, Siegel, Simon, Sokka, Strand, Tyndall, Verstappen, Wolfe, Zink, Boers.

Acquisition of data. Felson, Smolen, Zhang, van Tuyl, Allaart, Emery, Furst, Gomez-Reino, White, Boers.

Analysis and interpretation of data. Felson, Smolen, Wells, Zhang, van Tuyl, Funovits, Aletaha, Allaart, Bathon, Bombardieri, Choi, Combe, de Wit, Dougados, Gomez-Reino, Keystone, Khanna, Kvien, Landewé, Michaud, Martin-Mola, Montie, Pincus, Richards, Siegel, Simon, Strand, Tugwell, van der Heijde, White, Wolfe, Boers.

REFERENCES

- Emery P, Salmon M. Early rheumatoid arthritis: time to aim for remission? *Ann Rheum Dis* 1995;54:944–7.
- Van Tuyl LH, Vlad SC, Felson DT, Wells G, Boers M. 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.
- Smolen JS, Aletaha D. What should be our treatment goal in rheumatoid arthritis today? *Clin Exp Rheumatol* 2006;24:S-7-13.
- Mottonen T, Hannonen P, Leirisalo-Repo M, Nissila M, Kautiainen H, Korpela M, et al, for the FIN-RACo Trial Group. Comparison of combination therapy with single-drug therapy in early rheumatoid arthritis: a randomised trial. *Lancet* 1999;353:1568–73.
- Emery P, Breedveld FC, Hall S, Durez P, Chang DJ, Robertson D, et al. Comparison of methotrexate monotherapy with a combination of methotrexate and etanercept in active, early, moderate to severe rheumatoid arthritis (COMET): a randomised, double-blind, parallel treatment trial. *Lancet* 2008;372:375–82.
- Burmester GR, Ferraccioli G, Flipo RM, Monteagudo-Saez I, Unnebrink K, Kary S, et al. Clinical remission and/or minimal disease activity in patients receiving adalimumab treatment in a multinational, open-label, twelve-week study. *Arthritis Rheum* 2007;59:32–41.
- Verstappen SM, Jacobs JW, van der Veen MJ, Heurkens AH, Schenk Y, ter Borg EJ, et al, for the Utrecht Rheumatoid Arthritis Cohort study group. Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial). *Ann Rheum Dis* 2007;66:1443–9.
- Quinn MA, Conaghan PG, O'Connor PJ, Karim Z, Greenstein A, Brown A, et al. Very early treatment with infliximab in addition to methotrexate in early, poor-prognosis rheumatoid arthritis reduces magnetic resonance imaging evidence of synovitis and damage, with sustained benefit after infliximab withdrawal: results from a twelve-month randomized, double-blind, placebo-controlled trial. *Arthritis Rheum* 2005;52:27–35.
- Combe B, Landewe R, Lukas C, Bolosiu HD, Breedveld FC, Dougados M, et al. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCSIT). *Ann Rheum Dis* 2006;66:34–45.
- Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al, for the T2T Expert Committee. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;69:631–7.
- Khanna D, Oh M, Furst DE, Ranganath V, Gold RH, Sharp JT, et al, for the Western Consortium of Practicing Rheumatologists. Evaluation of the preliminary definitions of minimal disease activity and remission in an early seropositive rheumatoid arthritis cohort. *Arthritis Rheum* 2007;57:440–7.
- Listing J, Strangfeld A, Rau R, Kekow J, Gromnica-Ihle E, Klopsch T, et al. Clinical and functional remission: even though biologics are superior to conventional DMARDs overall success rates remain low—results from RABBIT, the German biologics register. *Arthritis Res Ther* 2006;8:R66.
- Mierau M, Schoels M, Gonda G, Fuchs J, Aletaha D, Smolen JS. Assessing remission in clinical practice. *Rheumatology (Oxford)* 2007;46:975–9.
- Sokka T, Hetland ML, Mäkinen H, Kautiainen H, Horslev-Petersen K, Luukkainen RK, et al, for the Questionnaires in Standard Monitoring of Patients with Rheumatoid Arthritis Group. Remission and rheumatoid arthritis: data on patients receiving usual care in twenty-four countries. *Arthritis Rheum* 2008;58:2642–51.
- Gossec L, Dougados M, Goupille P, Cantagrel A, Sibilia J, Meyer O, et al. Prognostic factors for remission in early rheumatoid arthritis: a multiparameter prospective study. *Ann Rheum Dis* 2004;63:675–80.
- Pinals RS, Masi AT, Larsen RA, and the Subcommittee of Criteria of Remission in Rheumatoid Arthritis of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for clinical remission in rheumatoid arthritis. *Arthritis Rheum* 1981;24:1308–15.
- Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Suppl* 1994;41:86–9.
- Fransen J, Creemers MC, van Riel PL. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology (Oxford)* 2004;43:1252–5.
- Aletaha D, Ward MM, Machold KP, Nell VP, Stamm T, Smolen JS. Remission and active disease in rheumatoid arthritis: defining criteria for disease activity states. *Arthritis Rheum* 2005;52:2625–36.
- Prevoe ML, van Gestel AM, van 't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Remission in a prospective study of patients with rheumatoid arthritis: American Rheumatism Association preliminary remission criteria in relation to the disease activity score. *Br J Rheumatol* 1996;35:1101–5.
- Prevoe ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts: development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;38:44–8.
- Van der Heijde D, Klareskog L, Boers M, Landewe R, Codreanu C, Bolosiu HD, et al. Comparison of different definitions to classify remission and sustained remission: 1 year TEMPO results. *Ann Rheum Dis* 2005;64:1582–7.
- Mäkinen H, Kautiainen H, Hannonen P, Sokka T. Is DAS28 an appropriate tool to assess remission in rheumatoid arthritis? *Ann Rheum Dis* 2005;64:1410–3.
- Landewe R, van der Heijde D, van der Linden S, Boers M. Twenty-eight-joint counts invalidate the DAS28 remission definition owing to the omission of the lower extremity joints: a comparison with the original DAS remission. *Ann Rheum Dis* 2006;65:637–41.
- Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;38:727–35.
- Emery P, Keystone E, Tony HP, Cantagrel A, van Vollenhoven R, Sanchez A, et al. IL-6 receptor inhibition with tocilizumab improves treatment outcomes in patients with rheumatoid arthritis refractory to anti-tumour necrosis factor biologicals: results from a 24-week multicentre randomised placebo-controlled trial [published erratum appears in *Ann Rheum Dis* 2009;68:296]. *Ann Rheum Dis* 2008;67:1516–23.
- Smolen JS, Kay J, Doyle MK, Landewe R, Matteson EL,

- Wollenhaupt J, et al. Golimumab in patients with active rheumatoid arthritis after treatment with tumour necrosis factor α inhibitors (GO-AFTER study): a multicentre, randomised, double-blind, placebo-controlled, phase III trial. *Lancet* 2009; 374:210–21.
28. Smolen JS, Beaulieu A, Rubbert-Roth A, Ramos-Remus C, Rovensky J, Alecock E, et al. Effect of interleukin-6 receptor inhibition with tocilizumab in patients with rheumatoid arthritis (OPTION study): a double-blind, placebo-controlled, randomised trial. *Lancet* 2008;371:987–97.
 29. Wells G, Boers M, Shea B, Brooks PM, Simon LS, Strand CV, et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;32:2016–24.
 30. Van Tuyl LH, Felson DT, Wells G, Smolen J, Zhang B, Boers M, for the American College of Rheumatology and the European League Against Rheumatism Committee to Define Remission for Clinical Trials. Evidence for predictive validity of remission on long-term outcome in rheumatoid arthritis: a systematic review. *Arthritis Care Res (Hoboken)* 2010;62:108–17.
 31. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. *Trials* 2007;8:38.
 32. 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. *Arthritis Rheum* 2008;59:1371–7.
 33. 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.
 34. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998;25: 198–9.
 35. American College of Rheumatology Committee to Reevaluate Improvement Criteria. A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. *Arthritis Rheum* 2007;57:193–202.
 36. Hewlett S, Wit M, Richards P, Quest E, Hughes R, Heiberg T, et al. Patients and professionals as research partners: challenges, practicalities, and benefits. *Arthritis Rheum* 2006;55:676–80.
 37. Bathon JM, Martin RW, Fleischmann RM, Tesser JR, Schiff MH, Keystone EC, et al. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med* 2000; 343:1586–93.
 38. Keystone EC, Kavanaugh AF, Sharp JT, Tannenbaum H, Hua Y, Teoh LS, et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. *Arthritis Rheum* 2004;50:1400–11.
 39. Weinblatt ME, Keystone EC, Furst DE, Moreland LW, Weisman MH, Birbara CA, et al. Adalimumab, a fully human anti-tumor necrosis factor α monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. *Arthritis Rheum* 2003;48:35–45.
 40. Klareskog L, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, et al, for the TEMPO (Trial of Etanercept and Methotrexate with Radiographic Patient Outcomes) Study Investigators. Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomised controlled trial. *Lancet* 2004;363:675–81.
 41. Breedveld FC, Weisman MH, Kavanaugh AF, Cohen SB, Pavelka K, van Vollenhoven R, et al, for the PREMIER Investigators. The PREMIER study: a multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;54:26–37.
 42. Landewe RB, Boers M, Verhoeven AC, Westhovens R, van de Laar MA, Markusse HM, et al. COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. *Arthritis Rheum* 2002;46:347–56.
 43. Smolen JS, Aletaha D, Grisar JC, Stamm TA, Sharp JT. Estimation of a numerical value for joint damage-related physical disability in rheumatoid arthritis clinical trials. *Ann Rheum Dis* 2010;69: 1058–64.
 44. Sharp JT, Lidsky MD, Collins LC, Moreland J. Methods of scoring the progression of radiologic changes in rheumatoid arthritis: correlation of radiologic, clinical and laboratory abnormalities. *Arthritis Rheum* 1971;14:706–20.
 45. Van der Heijde DM, van Riel PL, Nuvér-Zwart IH, Gribnau FW, van de Putte LB. Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. *Lancet* 1989;1:1036–8.
 46. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137–45.
 47. Landewe R, van der Heijde D. Radiographic progression depicted by probability plots: presenting data with optimal use of individual values. *Arthritis Rheum* 2004;50:699–706.
 48. Smolen JS, Han C, Bala M, Maini RN, Kalden JR, van der Heijde D, et al, for the ATTRACT Study Group. Evidence of radiographic benefit of infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis With Concomitant Therapy study. *Arthritis Rheum* 2005;52:1020–30.
 49. Smolen JS, Han C, van der Heijde DM, Emery P, Bathon JM, Keystone E, et al, for the Active-Controlled Study of Patients Receiving Infliximab for the Treatment of Rheumatoid Arthritis of Early Onset (ASPIRE) Study Group. Radiographic changes in rheumatoid arthritis patients attaining different disease activity states with methotrexate monotherapy and infliximab plus methotrexate: the impacts of remission and tumor necrosis factor blockade. *Ann Rheum Dis* 2009;68:823–7.
 50. Landewe R, van der Heijde D, Klareskog L, van Vollenhoven R, Fatenejad S. Disconnect between inflammation and joint destruction after treatment with etanercept plus methotrexate: results from the trial of etanercept and methotrexate with radiographic and patient outcomes. *Arthritis Rheum* 2006;54:3119–25.
 51. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G, et al. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford)* 2003;42:244–57.
 52. Aletaha D, Smolen JS. The definition and measurement of disease modification in inflammatory rheumatic diseases. *Rheum Dis Clin North Am* 2006;32:9–44.
 53. Kapral T, Dernoschnig F, Machold KP, Stamm T, Schoels M, Smolen JS, et al. Remission by composite scores in rheumatoid arthritis: are ankles and feet important? *Arthritis Res Ther* 2007;9:R72.
 54. Linde L, Sorensen J, Ostergaard M, Horslev-Petersen K, Hetland ML. Does clinical remission lead to normalization of EQ-5D in patients with rheumatoid arthritis and is selection of remission criteria important? *J Rheumatol* 2010;37:285–90.
 55. Balsa A, de Miguel E, Castillo C, Peiteado D, Martin-Mola E. Superiority of SDAI over DAS-28 in assessment of remission in rheumatoid arthritis patients using power Doppler ultrasonography as a gold standard. *Rheumatology (Oxford)* 2010;49:683–90.
 56. Fuchs HA, Pincus T. Reduced joint counts in controlled clinical trials in rheumatoid arthritis. *Arthritis Rheum* 1994;37:470–5.
 57. Smolen JS, Breedveld FC, Eberl G, Jones I, Leeming M, Wylie

- GL, et al. Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum* 1995;38:38–43.
58. Vita AJ, Terry RB, Hubert HB, Fries JF. Aging, health risks, and cumulative disability. *N Engl J Med* 1998;338:1035–41.
 59. Aletaha D, Strand V, Smolen JS, Ward MM. 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.
 60. Dawes PT, Fowler PD, Clarke S, Fisher J, Lawton A, Shadforth MF. Rheumatoid arthritis: treatment which controls the C-reactive protein and erythrocyte sedimentation rate reduces radiological progression. *Br J Rheumatol* 1986;25:44–9.
 61. Van Leeuwen MA, van der Heijde DM, van Rijswijk MH, Houtman PM, van Riel PL, van de Putte LB, et al. Interrelationship of outcome measures and process variables in early rheumatoid arthritis: a comparison of radiologic damage, physical disability, joint counts, and acute phase reactants. *J Rheumatol* 1994;21:425–9.
 62. Van der Heijde DM, van Riel PL, van Leeuwen MA, van 't Hof MA, van Rijswijk MH, van de Putte LB. Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis: a prospective follow-up study of 147 patients. *Br J Rheumatol* 1992;31:519–25.
 63. Pincus T, Yazici Y, Bergman M, Maclean R, Harrington T. A proposed continuous quality improvement approach to assessment and management of patients with rheumatoid arthritis without formal joint counts, based on quantitative Routine Assessment of Patient Index Data (RAPID) scores on a Multidimensional Health Assessment Questionnaire (MDHAQ). *Best Pract Res Clin Rheumatol* 2007;21:789–804.
 64. Pincus T, Chung C, Segurado OG, Amara I, Koch GG. An index of patient reported outcomes (PRO-Index) discriminates effectively between active and control treatment in 4 clinical trials of adalimumab in rheumatoid arthritis. *J Rheumatol* 2006;33:2146–52.
 65. Kirwan JR, Hewlett S. Patient perspective: reasons and methods for measuring fatigue in rheumatoid arthritis. *J Rheumatol* 2007;34:1171–3.
 66. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. *Arthritis Rheum* 2006;54:3761–73.
 67. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008;58:2958–67.
 68. Wolfe F. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. *J Rheumatol* 1997;24:1477–85.