

---

# ACR remission criteria and response criteria

---

V. K. Ranganath<sup>1</sup>, D. Khanna<sup>2</sup>, H. E. Paulus<sup>1</sup>

---

Veena K. Ranganath, Dinesh Khanna,  
Harold E. Paulus.

<sup>1</sup>Division of Rheumatology, Department of Medicine, University of California, Los Angeles, California; <sup>2</sup>Division of Immunology, Department of Internal Medicine; Institute for Study of Health, University of Cincinnati, Ohio, and Veterans Affairs Medical Center, Cincinnati, Ohio, USA.

Supported in part by grants from ASP/REF/ACR Junior Career Development Award in Geriatrics, UCLA Claude Pepper Older Americans Independence Center (OAIC) Career Development Award, Arthritis and Scleroderma Foundations, and National Institutes of Health BIRCWH award.

Please address correspondence to: Veena K. Ranganath, MD, UCLA, Department of Medicine, Division of Rheumatology, Rehabilitation center, Room 32-59, 1000 Veteran Avenue, Los Angeles, CA 90095-1670, USA.

E-mail: vranganath@mednet.ucla.edu

*Clin Exp Rheumatol* 2006; 24 (Suppl. 43): S14-S21.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2006.

**Key words:** Remission, rheumatoid arthritis, outcome measures, historical.

## ABSTRACT

*As additional DMARDs have been added to the armamentarium of rheumatologists over the last 60 years, the approach to the treatment of rheumatoid arthritis has changed. Many clinical studies now are geared toward evaluating the concept of eradicating inflammation as a method to seek the elusive goal of sustained remission in RA. One of the first descriptions of remission in 'RA' was by Short et al in 1948, when he documented the natural progression of the disease. Since that time, various criteria have been developed to define RA remission utilizing clinical, radiographic, and laboratory measures. The most stringent of criteria is the American College of Rheumatology Remission Criteria, developed in 1980, which consists of clinical symptoms and signs of inflammation including fatigue, joint pain, morning stiffness, joint tenderness, joint swelling, and erythrocyte sedimentation rate (ESR). Several reports have compared ACR remission criteria to Disease Activity Score (DAS) values to identify equivalent DAS remission values, and these have been extrapolated to modified versions of the DAS, the Simple Disease Activity Index (SDAI), and Clinical Disease Activity Index (CDAI). The ACR remission criteria and the response measures were not designed for use as the target or goal for the clinical management of individual RA patients in routine clinical practice. Nevertheless, rheumatologists yearn for the eradication of inflammation in all RA patients, and attaining remission may be achievable in the future.*

## Introduction

Despite the significant advances in disease-modifying therapy and clinical trial methodology, it is unusual for a physician to have the pleasure of telling a patient with rheumatoid arthritis (RA) that his or her unrelenting disease is in remission. As seen with disease activity measures, "remission" in RA is

difficult to define precisely with a single disease activity measure. Clinical remission is defined as an absence of joint inflammation and extraarticular disease activity (1) and is now being considered a target for RA therapy. The approach to the treatment of RA has changed over the past 60 years as additional disease-modifying antirheumatic drugs (DMARDs) with good efficacy and toxicity profiles have been added to the armamentarium of rheumatologists. Many studies now are geared toward evaluating the concept of eradicating inflammation as close as possible to symptom onset, as a method to seek the elusive goal of remission in RA.

One of the first descriptions of remission in RA was by Short *et al.* (2, 3), in which he illustrated the natural history of the disease over a 24-year period. In 1948, Short and Bauer described a series of 300 RA patients (above the age of 12) admitted to Massachusetts General Hospital between 1930 and 1936. Approximately 38 patients with peripheral arthritis were found to have rheumatoid spondylitis, but were included in the cohort, while patients with a clearly erroneous diagnosis were excluded. Patients' hospital stay was usually 3 to 4 weeks and included the following therapeutic interventions: rest, analgesics, exercises, heat application to joints, diet with supplementary vitamins, fever therapy, blood transfusions, treatment of infections, and orthopedic procedures. Of the 293 patients, 38 patients with more severe disease required readmission. Patients were seen in clinic at least once or twice a year, however some patients were seen at irregular intervals. During the 12 to 18 years of observation, 56 patients died of "causes unrelated to the arthritis". Short *et al.* classified patients as being in RA remission "if the disease was inactive, the patients were asymptomatic and examination of the joints was negative except for residual deformity". Patients were assessed in 1937,

1947, and 1954. Subanalyses were performed excluding rheumatoid spondylitis patients. In 1937 approximately 17% of patients met their remission criteria, and the average duration of remission was 21 months. When assessed in 1947, 50% of patients who were in remission in 1937 were still in remission, and 17.4% of the patients remaining under study were noted to be in remission. Patients above the age of 40 were less likely to be in remission compared to patients below the age of 40 (8.5% versus 21%). In addition, patients with disease duration of greater than 1 year at study entry were noted to have a remission rate of 4.7%, compared to 37% of patients with less than 1 year of disease duration. In 1954 only 13% of the remaining 174 patients met the authors' criteria for remission. This study pioneered analysis of remission in RA and initiated the process of establishing remission criteria as a goal for RA treatment.

In 1980, a Subcommittee for Criteria of Remission in Rheumatoid Arthritis of the American Rheumatism Association (ARA) Committee on Diagnostic and Therapeutic Criteria was convened to develop criteria for clinical remission in RA. Complete remission was defined as "total absence of all articular and extraarticular inflammation and immunological activities related to RA" (1). To achieve this, 35 practicing rheumatologists were asked to provide data on their patients using a RA data collection form and to classify the patients' disease activity into four categories: complete remission without treatment, complete remission with treatment, partial remission, and active disease. They recorded demographic information, past and present symptoms (including extra-articular manifestations), results of a joint examination, laboratory data, and radiographic data. Through statistical methods, each variable was assessed for the strength of its capacity to discriminate between patients in complete remission versus those in partial remission or active disease.

The clinical study included 344 patients; 63 patients were considered to be in complete remission without treatment,

112 in remission with treatment, 93 in partial remission, and 76 to have active disease. The 175 patients in complete clinical remission, with and without treatment for RA, did not differ significantly and were combined for the analyses. The presence of rheumatoid factor positivity was slightly higher in the active disease group (87%) when compared to patients in the complete remission group (74%). Some patients in complete remission had rheumatoid nodules and Sjögren's syndrome. Radiographs at a single time point were used as an index of disease severity, and no new radiographs were required in the study. Morning stiffness was a discriminating variable when comparing patients in complete remission (18%) to those with active disease (96%). It was even more discriminating when evaluating patients' duration of morning stiffness with the cutoff of 15 minutes (p-value < 0.01).

The Subcommittee then proposed the following definition for ARA (now American College of Rheumatology [ACR]) complete clinical remission that had the highest face and discriminatory validity. The patient should meet five of the following six criteria for at least 2 consecutive months: morning stiffness  $\leq$  15 min, no fatigue, no joint pain (by history), no swollen joints, no tender joints, and erythrocyte sedimentation rate (ESR) < 30 mm/h for female or 20 mm/h for male. The sensitivity and specificity for meeting five out of six criteria for 2 consecutive months were 72% and 92% compared to patients with partial remission, and 72% and 100% compared to patients with active disease. A period of 2 months' time required to be in remission was arbitrarily chosen, and 90% of the patient population met this criterion. Multiple calculations of 2:1 or 3:1 weighing of variables with better discriminating ability produced the same results as using equal weights of the six variables.

The Outcome Measures in Rheumatology (OMERACT) group recently proposed preliminary definitions of minimal disease activity (MDA) (4). This was a concept that Pinals *et al.* considered in their discussion of

remission criteria in RA; MDA might characterise a population of patients in "transition" to reaching remission. MDA was defined as a disease state that is "between high disease state and remission", and was developed to fill the need for a measurable disease activity goal that could be attained in clinical practice, since complete clinical remission in RA is a rarity. A patient in remission would also meet MDA criteria. Aletaha and Smolen have proposed definitions for MDA and remission (5) based on the Disease Activity Score 28 (DAS28), Simplified Disease Activity Index (SDAI), and Clinical Disease Activity Index (CDAI) (6).

Our group recently assessed different definitions of clinical remission and MDA at 6, 12, and 24 months in a well characterised cohort of early seropositive, DMARD-naïve, RA patients (n = 200) with active disease at entry, treated with traditional DMARDs in the prebiologic era (7) (in press). At baseline, none of the 200 patients were in MDA or remission. We modified the ACR remission criteria by assessing patients cross-sectionally—that is, at single points in time, rather than over a consecutive 2-month period. The modified ACR remission definition was the most stringent 0.7%, 0%, and 0% were in clinical remission compared to 3% to 15%, 4% to 24%, and 6% to 33% at 6, 12, and 24 months, respectively, for other published criteria for clinical remission (5) (see chapters "DAS remission cut points" by Fransen *et al.*; "Definitions of remission for rheumatoid arthritis and review of selected clinical cohorts and randomised clinical trials for the rate of remission" by Mäkinen *et al.*). Depending on the MDA definition used, 20% to 32%, 27% to 32%, and 30% to 48% of patients were in MDA at 6, 12, and 24 months, respectively. Patients who achieved either MDA or remission had lower Health Assessment Questionnaire-Disability Index (HAQ-DI) and radiographic scores compared to patients who did not achieve either.

#### Status and response measures in RA

It is important to understand the concepts of "status" and "response"

measures, which are used to describe disease activity in RA. "Status" measures assess disease activity at a specific point in time, and "response" measures assess how disease activity changes over time, for example, response to medication (8). Remission in RA is considered to be a status measure. Other status measures include MDA measures, DAS, and its variations, HAQ-DI, and Sharp Score of radiographic evidence of joint damage. These measures are important when the practicing rheumatologist evaluates treatment strategies in individual RA patients. Response measures evaluate change in clinical status over time in clinical trials to determine efficacy but also can be implemented in longitudinal observational studies to evaluate clinical change over time. Two types of response measures in current usage are the American College of Rheumatology 20%, 50%, and 70% improvement (ACR 20/50/70) criteria (9), and the European League of Associations for Rheumatology (EULAR) Improvement Criteria (10) and its variations. Three domains can be used to describe the long- and short-term consequences of disease activity in RA and should be considered in defining remission: clinical signs and symptoms of inflammation, functional impairment, and structural joint damage (Table I). These domains will be discussed in the following section in more detail.

### Importance of clinical trial design and metrology in DMARD development

The treatment of RA has been advanced by quantitative measurement of RA manifestations (metrology) and in clinical trial methodology. It is difficult to deem a therapeutic agent as beneficial in a disease such as RA, in which patients suffer a progressive, chronic inflammatory disease with spontaneous, robust flares and other times of quiescence; although the immediate, dramatic benefit of cortisone was rapidly apparent to everyone and did not require formal clinical trends. Clinical manifestations are not always uniform in presentation of signs and symptoms, and the beneficial effect of

**Table I.** Assessing the consequences of rheumatoid arthritis.

Clinical signs and symptoms of inflammation	Functional disability	Joint damage
<ul style="list-style-type: none"> <li>• Swollen joint count</li> <li>• Tender joint count</li> <li>• Morning stiffness</li> <li>• Acute phase reactants (ESR, CRP)</li> </ul>	<ul style="list-style-type: none"> <li>• HAQ-DI, mHAQ, SF-36</li> <li>• ARA functional class</li> <li>• Walking time</li> <li>• Grip strength</li> <li>• Work status</li> </ul>	<ul style="list-style-type: none"> <li>• Clinical joint deformity</li> <li>• Sharp scores, Larsen scores, joint space narrowing, erosions, malalignment</li> <li>• MRI and ultrasound of hands</li> <li>• Joint replacement surgery</li> </ul>
Fatigue Co-morbidities		
Pain visual analog scale Range of motion Physician global assessment Patient global assessment		

\*Adapted from Pincus T, Sokka T, 2005 (52).

the treatment often takes months to appreciate. For example, the beneficial effect of parenteral gold in RA was debated for 30 years, until it was finally proven in 1960 by the Empire Rheumatism Council's controlled clinical trial (11). Our rheumatologist predecessors who routinely treated patients with chronic, debilitating RA, identified these complicated issues and were among the first to employ the concepts of randomisation, blinding, and use of a control group in clinical trials. These shrewd clinician researchers were some of the first to pursue a broad spectrum of objective, semi-objective, and subjective methods to quantify patient status and changes in status over time.

Appropriate classification and characterisation of patients with RA was the first hurdle to tackle. Short, Bauer, and Reynolds reviewed the history of the "numerical method in the study of disease" in the monograph "Rheumatoid Arthritis" (Harvard University Press, 1957) (3). The detailed tabulation of characteristics in RA patients and numerical basis for diagnostic categorisation quantified the previous method of diagnosis by expert opinion. In 1945, Fletcher and Lewis-Faning published a detailed clinical and statistical account of 1,000 cases of chronic rheumatism, including 254 patients with RA (12, 13). They emphasised clinical and radiographic criteria to help distinguish RA from degenerative joint disease. Lewis-Faning later presented a meticulous statistical analysis

for the Empire Rheumatism Council in 1950, of 532 RA patients matched for age, sex, and "civil state" (i.e., socioeconomic status) with non-RA controls to determine factors related to the onset of RA (14).

In 1955, the Empire Rheumatism Council also published "Multicenter controlled trial comparing cortisone acetate and acetylsalicylic acid in long-term treatment of rheumatoid arthritis" (15). The same group refined their methodology with a series of multicenter controlled clinical trials of corticosteroids, gold, and high dose aspirin (11, 16, 17). "A controlled study of chloroquine as an antirheumatic agent" was published in 1958 (18), and "Chloroquine in rheumatoid arthritis: a double-blind fold trial of treatment for one year" was published in 1960 (19).

These studies included detailed descriptions of which joints were involved (in what percent of patients), as well as joint swelling, limitation of motion, effusion, heat, redness, fatigue, and osseous overgrowth, although formal joint counts were not done. Nodules, fever, lymphadenopathy, anemia, and stiffness were also tabulated. During the 1950s and 1960s, various methods to describe numerical quantitation of the magnitude of individual signs and symptoms of RA were evaluated, including Likert (20) and visual analogue scales of pain (21) and global well-being, durations of morning stiffness and fatigue (22), circumference (ring size) of swollen joints (23), grip strength (24), pain on pressure (dolori-

meter) (25), 50 feet walk-time (26), and measurement of joint motion with a goniometer (27).

One of the earliest indices to evaluate disease activity was Steinbrocker's Therapeutic Scorecard in 1946, which included an arbitrary set of clinical signs and symptoms of inflammation with seven out of nine of the items falling into this domain (joint swelling, joint motion, joint tenderness, ESR, haemoglobin, pain, well-being), with the other two items dealing with functional capacity and weight (28). The scorecard was filled out by using a debit system. Thus, if patients described any of the above symptoms, a specific percentage was deducted from a total theoretical score of 100 percent, representing a healthy patient. The scorecard heavily emphasised clinical signs and symptoms of RA; only 5% of the total was allotted to functional status, and structural joint damage was not directly included in the scoring system. Similarly, Lansbury's Systemic Manifestations of Rheumatoid Activity in 1954 also emphasised clinical signs and symptoms of inflammation. Lansbury's index included ESR, pain on motion, muscle weakness, morning stiffness, fatigue, anemia, pain at rest, and fever (29).

A few years later in 1949, Steinbrocker proposed four-point global scales to quantitate functional status/disability (30), and Lansbury proposed a formal joint count, weighted by joint size in 1958 (31). Many other composite measures have been proposed to evaluate disease activity in RA, including the pooled index (32), the discriminant analysis (33), the Cooperative Systematic Studies of the Rheumatic Diseases (CSSRD) joint count (34), and others (22). These indices did not gain widespread use, though interestingly Steinbrocker's Therapeutic Scorecard included all of the items later selected for the ACR Core Set (35). The four-point global clinical and radiographic scales were used by rheumatologists, although they were not sufficiently sensitive to discriminate modest treatment effects.

The Cooperating Clinics Committee of the ARA, statistically guided by Donald

Mainland, evaluated the statistical stability and characteristics of numerous outcome measures in "A seven-day variability study of 499 patients with peripheral RA" in 1965, and proposed the widely used unweighted joint counts of 66 swollen and 68 tender joints (36). This "ARA-Index" 66/68 joint count was further described and defined by Mainland in 1967, and was considered to measure RA activity, along with morning stiffness, grip strength, and ESR (37). In 1968 the Ritchie articular index (RAI) of tender joints was published, in which some joints are grouped, the distal interphalangeal joints are omitted, and the degree of tenderness is graded from 0 to 3 for each joint and joint group (38). Meanwhile, the nonsteroidal anti-inflammatory drug (NSAID) indomethacin had been identified by screening numerous compounds for their ability to rapidly decrease the acute swelling induced by the injection of carrageenan into a rat paw (39). This rapid, relatively inexpensive screening method was widely applied during the late 1960s and 1970s to identify a large number of commercially viable NSAIDs. However, statistically valid trial evidence of clinical benefit was required before regulatory approval for marketing could be obtained. Multiple outcome measures and clinical trial designs were used by various sponsors.

In 1971 an advisory committee to the US Food and Drug Administration (FDA) published the first "Guidelines for clinical evaluation of nonsteroidal anti-inflammatory drugs" (40), which enhanced and expedited the development and marketing of a large number of NSAIDs. These guidelines, which were heavily influenced by findings from the 7-day variability study (36), helped to standardise industry-sponsored clinical trials. Under the watchful prodding of the FDA rheumatologist John Harter, MD, clinical data submitted in support of new drug applications were used to test and improve these guidelines, which were expanded to include DMARDs and later biologic agents for RA. Furthermore, principles of these guidelines were then expanded to

include the assessment of drugs for other rheumatic diseases. Additional outcome measures of function [e.g., HAQ (41), Arthritis Impact Measurement Scales (AIMS) (42)], and radiographic damage [e.g., Sharp score (43) and Larsen score (44)] were added to "signs and symptoms" as potential indications for antirheumatic agents.

Analysis of clinical trials remained difficult however, because of the large number of outcome measures being used. When analysed within the same trial, improvement of some measures frequently was contradicted by worsening in other measures, and sponsors could emphasise the measures that improved while ignoring those that did not improve. In several papers in 1988 and 1989 Dixon *et al.* (45) and Anderson *et al.* (46) analysed the relative sensitivity to change of various outcome measures and discussed which measures were most efficient for RA clinical trials. General interest of clinical trialists led to an international conference with agreement on a core set of measures recommended for all RA clinical trials (35).

Traditional clinical assessments of RA attempted to measure the cardinal signs of inflammation: joint pain, joint tenderness, joint swelling, range of motion, joint circumference, grip strength, walking time, morning stiffness, and ESR. Published reports from controlled clinical trials of NSAIDs and DMARDs focused on the statistical differences of single clinical assessments of RA inflammation in order to differentiate between one therapeutic intervention and another, and from placebo. It was understood that these clinical assessments were interrelated, but were independently analysed. If all clinical assessments of RA were statistically superior, then the treatment in question was considered more effective than its comparator; however there was difficulty in identifying the superior therapy when only some assessments were statistically superior. In addition, there was a lack of a generally accepted method for estimating improvement in an individual RA patient in clinical trials.

The problem of discordant outcome

**Table II.** Components of various response and remission criteria.

Paulus improvement criteria (47)	ACR20 improvement criteria (9)	ACR remission criteria (1)
20% improvement in 4 of the 6 parameters	An improvement to the 20%, 50%, or 70% level in the parameters outlined, require improvement in both SJC and TJC, and 3/5 other items	≥ 5 of the following present at least 2 consecutive months
Swollen joint count	Swollen joint count	Morning stiffness ≤ 15 minutes
Tender joint count	Tender joint count	No fatigue
Morning stiffness	Pain visual analog scale	No joint pain
Patient global assessment	Patient global assessment	No joint tenderness or pain on motion
Physician global assessment	Physician global assessment	No soft tissue swelling in joints or tendon sheaths
ESR	ESR or CRP	ESR (Westergren method) ≤ 30 mm/h for a female or 20 mm/h for a male
	Functional questionnaire	

measures within the same clinical trial prompted a search for a single composite measure [e.g., pooled index (32), the discriminant analysis (33), the CSSRD joint count (34)] that could be used to evaluate each subject in a clinical trial as improved or not improved. In an effort to assess clinical signs and symptoms of inflammation as a single measure, Paulus *et al.* with the CSSRD Group proposed a method to evaluate them collectively, now called the “Paulus Improvement Criteria” (Table II). The CSSRD group, a National Institutes of Health sponsored program led by John Ward and the University of Utah, had conducted four placebo-controlled clinical trials with standard DMARDs. Collectively, a total of 198 patients who were randomised to placebo and drug met eligibility criteria. Of the 198 patients, 27 withdrew due to lack of efficacy. The standard measures of efficacy included in each

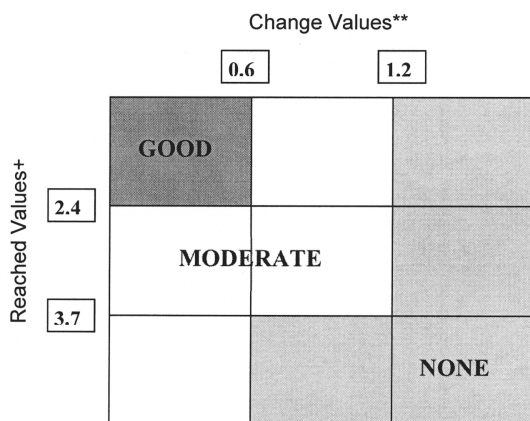
of the trials were: morning stiffness, CSSRD joint pain/tenderness score, CSSRD joint swelling score, patient’s overall assessment of current disease activity, physician’s overall assessment of current disease severity, and ESR. These 6 measures were arbitrarily selected to develop the Paulus Criteria. By using the pooled data of the placebo-treated patients in these DMARD trials, as well as the data for the active DMARD-treated arms, Paulus, Egger, Williams, *et al.* (47) were able to formulate a composite measure of improvement (≥ 20% improvement in 4 of 6 measures that were used in all of the Cooperating Clinics trials), which clearly differentiated the proportion of patients improved during placebo (4%-10%) from those improved by the various DMARDs (16%-50%) (Table II). This method depicted relatively unambiguous outcomes to controlled clinical trials and decreased the number

of subjects needed to recognise an efficacious new DMARD.

In the late 1980s and early 1990s, recombinant technology had permitted the production of many potentially therapeutic biologic products, but sponsors were hesitant to test them as DMARDs in clinical trials because of the anticipated expense entailed by large prolonged clinical trials with ambiguous results. The proposed composite measure was successfully applied in the development of several early biologic agents (48). This move proved to be successful and led to an international conference of clinical trialists who agreed to the enhanced and improved ACR preliminary definition of improvement in RA (9); improvement criteria were derived from the core components and set (35). ACR20 improvement criteria that have become a de facto standard for DMARD trials (≥20% of improvement in tender joint count and swollen joint count plus ≥20% improvement in three of the following: patient pain, patient global disease activity, physician global disease activity, physical function, e.g., HAQ-DI and acute-phase reactants, ESR, or C-reactive protein [CRP]) (9).

The ACR20 has performed well in controlled clinical trials of new therapies (49) and was quickly expanded to include ACR50 and ACR70 improvement to recognise more effective therapies. Its major weakness is that it is tied to relative improvement compared to baseline and does not give much information about the absolute status of the subject’s RA. Patients with 20% improvement in mild RA are not really comparable to patients with 20% improvement in severe RA. For this reason ACR20/50/70 should be restricted to analysis of clinical trials or similar longitudinal outcome studies. It is not useful in the clinical management of individual patients.

At about the same time, van Riel, van der Heijde, and associates developed the DAS index (50, 51). Serial assessments of tender and swollen joint counts, ESR, and patient global assessment (GH = global health) were recorded for a panel of RA patients at times of poorly controlled RA (e.g.,



**Fig. 1.** EULAR response criteria (based on DAS44/ESR-4 item).  
 \*DAS44/ESR-4 item Score =  $0.53938 \sqrt{RAI} + 0.06465 \cdot SJC44 + 0.331 \ln(ESR) + 0.00722 \cdot GH$   
 \*\*Change values = Baseline DAS value minus current value.  
 \*Reached values = Current DAS score.

when a new DMARD was needed), and when well-controlled (e.g., no change in DMARD for 1 year or longer). These were compared statistically to arrive at a mathematical formula to express the degree of disease activity on an arbitrary scale at a given point in time. The domains of functional disability and joint damage are not included. Patients with similar DAS score have similar degrees of RA activity, and DAS scores change with improvement or worsening of RA activity. Standard values for change in score and current score have been established as the EULAR Improvement Criteria (Fig. 1) that is, none, moderate, or good improvement, that can be used in clinical trials as a composite outcome measure (10).

The ACR remission criteria strongly consider clinical signs and symptoms of inflammation with all six items belonging to the domain of clinical symptoms and signs of inflammation: fatigue, joint pain, morning stiffness, joint tenderness, joint swelling, and ESR. Several reports have compared ACR remission criteria to DAS values to identify equivalent DAS remission values. Both of these indices omit effects of RA on functional disability and structural joint damage (5). Two other disease activity measures are derived from the DAS, the Simple Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI); cut points of SDAI and CDAI have been developed to describe remission in RA (6).

Neither ACR20 nor EULAR improvement criteria adequately measure changes in physical function or changes in radiographic evidence of joint damage. These have been established as separate outcomes by the FDA and are generally included in DMARD development in order to obtain FDA-approved indications for improvement of physical function and for prevention of radiographic joint damage.

#### **Competing and complementary roles of stake-holders in development and use of remission and response criteria**

Early development of DMARDs involved efforts of rheumatologists to

adapt agents for treatment of RA from drugs that had been approved for other disease indications. Examples include gold salts, antimalarial drugs, penicillamine, azathioprine, cyclophosphamide, chlorambucil, methotrexate, and cyclosporine. The spontaneous development and off-label use of these drugs encouraged the FDA, through its Arthritis Advisory Committee, to develop and publish guidelines for clinical evaluation, first of NSAIDs, then of DMARDs and biologic agents, with the implicit assurance that new agents that satisfied the guidelines could be approved and marketed as DMARDs. As the scientific basis for DMARD activity was explored in basic science laboratories, pharmaceutical and biotechnology companies began to explore potential therapeutic applications. Around 1990, the development of a composite criterion for response based on the Cooperating Clinics DMARD studies, rapidly followed by the ACR core set of preferred outcome measures and international agreement on the ACR20 improvement criteria, set the stage for clinical development and eventual approval of leflunomide and biologic agents. The success of these agents has encouraged the development and assessment of many new agents with specific biologic targets, some of which are now approved for use in clinical practice.

ACR and EULAR have been active participants in this process. ACR and EULAR committees continue to compete and collaborate in the development, application, and refinement of response criteria and status measures. OMER-ACT has been a valuable forum for international collaboration in the development and application of standards for outcome measurement in a broad spectrum of rheumatic diseases by interested clinical trialists, statisticians, academicians, pharmaceutical industry scientists, and regulatory agencies.

A continual attempt to refine and improve the methods of clinical assessment remains in progress. Currently, the ACR has established a Quality Measures Committee with a subcommittee to standardise the ACR approval process for disease classification criteria

sets and for response criteria sets. Their policies are detailed in a recent editorial in *Arthritis and Rheumatism*, *Arthritis Care and Research* (8).

In essence, it is suggested that representative expert opinion be gathered by a defined Delphi method and refined by Nominal Group Techniques to develop potential consensus criteria. These candidate criteria should then be refined using appropriate cases and controls to determine sensitivity and specificity. Final validation requires a different set of cases, controls, and experts. There are three categories of ACR approval for a criteria set. Unendorsed "proposed" criteria may have been developed by expert consensus but have not yet been refined with appropriate cases, controls, and statistical validation. ACR endorsed "provisional" criteria have been validated with appropriate cases and controls and can be used in clinical research. Full endorsement of "final" criteria may occur after the "provisional" criteria have been used and validated in independent clinical trials and/or clinical studies and have been generally accepted by users. Of course, a criteria set is never truly "final", and one should anticipate periodic review, refinement, and revision of fully endorsed criteria sets.

#### **Use of remission and response criteria in clinical practice**

The ACR remission criteria and the response measures are neither designed for nor intended for use as the target or goal for the clinical management of individual RA patients in routine clinical practice at this point in time. If one wishes to treat a patient as intensively as necessary to achieve a certain absolute degree of disease suppression (e.g., "remission" or "minimal disease activity"), then the target will be the disease activity thresholds specified by that status measure. The relative value of adjusting treatment regimens to attain a specified disease remission threshold has not yet been determined but is a very different concept from the method used to describe ACR remission criteria as discussed in this paper. The ACR20 definition of improvement and

EULAR criteria are designed to encourage DMARD and biologic agent discovery and development, by providing clear “yes” or “no” improvement criteria when comparing a potential new agent with placebo. These criteria improve the development of new treatments of RA, but they are not properly applicable as targets for the clinical management of individual RA patients at this time. After gathering all available patient data and considering the available evidence from guidelines/literature, judicious use of clinical judgment remains necessary when deciding the best treatment option for the individual RA patient.

## References

- PINALS RS, MASI AT, LARSEN RA: Preliminary criteria for clinical remission in rheumatoid arthritis. *Arthritis Rheum* 1981; 24: 1308-15.
- SHORT CL, BAUER W: The course of rheumatoid arthritis in patients receiving simple medical and orthopedic measures. *N Engl J Med* 1948; 238: 142-8.
- SHORT CL, BAUER W, REYNOLDS WE: Rheumatoid Arthritis. Cambridge, Mass, Harvard University Press, 1957.
- WELLS GA, BOERS M, SHEA B *et al.*: Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005; 32: 2016-24.
- 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.
- ALETAHA D, SMOLEN J: The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005; 23 (5 Suppl. 39): S100-S108.
- KHANNA D, OH M, FURST DE *et al.*: Evaluation of the Preliminary Definitions of Minimal Disease Activity and Remission in an Early Seropositive Rheumatoid Arthritis Cohort. *Arthritis Rheum* 2006 (in press).
- SINGH JA, SOLOMON DH, DOUGADOS M *et al.*: Development of classification and response criteria for rheumatic diseases. *Arthritis Rheum* 2006; 55: 348-52.
- FELSON DT, ANDERSON JJ, BOERS M *et al.*: American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995; 38: 727-35.
- VAN GESTEL AM, PREVOO ML, VAN'T HOF MA, VAN RIJSWIJK MH, VAN DE PUTTE LB, VAN RIEL PL: Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996; 39: 34-40.
- EMPIRE RHEUMATISM COUNCIL: Gold therapy in rheumatoid arthritis: Final report of a multicentre controlled trial. *Ann Rheum Dis* 1961; 20: 315-34.
- FLETCHER E, LEWIS-FANING E: The chronic rheumatic diseases, with special reference to chronic arthritis; a survey based on 1,000 cases. *Post-Grad M J* 1945; 21: 1-13, 54-63.
- LEWIS-FANING E: A Statistical study of 1,000 cases of chronic rheumatism. *Post-Grad M J* 1945; 21: 137-46, 176-85.
- LEWIS-FANING E: Report on an enquiry into the aetiological factors associated with rheumatoid arthritis. *Ann Rheum Dis* 1950; 9: 1950.
- EMPIRE RHEUMATISM COUNCIL: Multi-centre controlled trial comparing cortisone acetate and acetyl salicylic acid in the long-term treatment of rheumatoid arthritis; results up to one year. *Ann Rheum Dis* 1955; 14: 353-70.
- EMPIRE RHEUMATISM COUNCIL: Multi-centre controlled trial comparing cortisone acetate and acetyl salicylic acid in the long-term treatment of rheumatoid arthritis; results of three years' treatment. *Ann Rheum Dis* 1957; 16: 277-89.
- EMPIRE RHEUMATISM COUNCIL: Gold therapy in rheumatoid arthritis. Report of a multicentre control trial. *Ann Rheum Dis* 1960; 19: 95-119.
- COHEN AS, CALKINS E: A controlled study of chloroquine as an antirheumatic agent. *Arthritis Rheum* 1958; 1: 297-312.
- FREEDMAN A, STEINBERG VL: Chloroquine in rheumatoid arthritis; a double blindfold trial of treatment for one year. *Ann Rheum Dis* 1960; 19: 243-50.
- LIKERT R: A technique for the measurement of attitudes. *Arch Psychol* 1932; 140: 44-60.
- HUSKISSON EC: Measurement of pain. *Lancet* 1974; 2: 1127-31.
- BUCHANAN WW, TUGWELL P: Traditional assessments of articular diseases. *Clin Rheum Dis* 1983; 9: 515-29.
- BOARDMAN PL, HART FD: Clinical measurement of the anti-inflammatory effects of salicylates in rheumatoid arthritis. *Br Med J* 1967; 4: 264-8.
- WRIGHT V: Some observations on diurnal variation of grip. *Clin Sci (Lond)* 1959; 18: 17-23.
- MCCARTHY DJ, JR., GATTER RA, PHELPS P: A dolorimeter for quantification of articular tenderness. *Arthritis Rheum* 1965; 8: 551-9.
- SPIEGEL JS, PAULUS HE, WARD NB, SPIEGEL TM, LEAKE B, KANE RL: What are we measuring? An examination of walk time and grip strength. *J Rheumatol* 1987; 14: 80-6.
- AMERICAN ACADEMY OF ORTHOPEDIC SURGEONS: Joint Motion: Method of Measuring and Recording. Edinburgh, Churchill Livingstone, 1966.
- STEINBROCKER O, BLAZER A: A therapeutic score card for rheumatoid arthritis: a standardized method of appraising results of treatment. *N Engl J Med* 1946; 235: 501-6.
- LANSBURY J: Quantitation of the activity of rheumatoid arthritis. I. A method for recording its systemic manifestations. *Am J Med Sci* 1954; 231: 616-21.
- STEINBROCKER O, TRAEGER CH, BATTERMAN RC: Therapeutic criteria in rheumatoid arthritis. *JAMA* 1949; 140: 659-62.
- LANSBURY J: Report of a three-year study on the systemic and articular indexes in rheumatoid arthritis; theoretic and clinical considerations. *Arthritis Rheum* 1958; 1: 505-22.
- SMYTHE HA, HELEWA A, GOLDSMITH CH: “Independent assessor” and “pooled index” as techniques for measuring treatment effects in rheumatoid arthritis. *J Rheumatol* 1977; 4: 144-52.
- MCGUIRE RJ, WRIGHT V: Statistical approach to indices of disease activity in rheumatoid arthritis. *Ann Rheum Dis* 1971; 30: 574-80.
- WILLIAMS HJ, WARD JR, READING JC *et al.*: Low-dose D-penicillamine therapy in rheumatoid arthritis. A controlled, double-blind clinical trial. *Arthritis Rheum* 1983; 26: 581-92.
- FELSON DT, ANDERSON JJ, BOERS M *et al.*: The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. *Arthritis Rheum* 1993; 36: 729-40.
- COOPERATING CLINICS COMMITTEE OF THE AMERICAN RHEUMATISMS ASSOCIATION: A Seven-Day Variability Study of 499 Patients With Peripheral Rheumatoid Arthritis. *Arthritis Rheum* 1965; 8: 302-34.
- MAINLAND D: The estimation of inflammatory activity in rheumatoid arthritis. Role of composite indices. *Arthritis Rheum* 1967; 10: 71-7.
- RITCHIE DM, BOYLE JA, MCINNES JM *et al.*: Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q J Med* 1968; 37: 393-406.
- RODNAN GP, BENEDEK TG: The early history of antirheumatic drugs. *Arthritis Rheum* 1970; 13: 145-65.
- McMILLEN J, LEER JA, RIDOLFO A *et al.*: Guidelines for clinical evaluation of nonsteroidal anti-inflammatory drugs. *Drug Res Rep* 1971; 15: S2-S26.
- FRIES JF, SPITZ P, KRAINES RG, HOLMAN HR: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
- MEENAN RF, GERTMAN PM, MASON JH: Measuring health status in arthritis. The arthritis impact measurement scales. *Arthritis Rheum* 1980; 23: 146-52.
- 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.
- LARSEN A: Radiological grading of rheumatoid arthritis. An interobserver study. *Scand J Rheumatol* 1973; 2: 136-8.
- DIXON JS, HAYES S, CONSTABLE PD, BIRD HA: What are the ‘best’ measurements for monitoring patients during short-term second-line therapy? *Br J Rheumatol* 1988; 27: 37-43.
- ANDERSON JJ, FELSON DT, MEENAN RF, WILLIAMS HJ: Which traditional measures should be used in rheumatoid arthritis clinical trials? *Arthritis Rheum* 1989; 32: 1093-9.
- PAULUS HE, EGGER MJ, WARD JR, WILLIAMS HJ: Analysis of improvement in individual rheumatoid arthritis patients treated with disease-modifying antirheumatic

- drugs, based on the findings in patients treated with placebo. The Cooperative Systematic Studies of Rheumatic Diseases Group. *Arthritis Rheum* 1990; 33: 477-84.
48. ELLIOTT MJ, MAINI RN, FELDMANN M *et al.*: Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor alpha (cA2) versus placebo in rheumatoid arthritis. *Lancet* 1994; 344: 1105-10.
49. MORELAND LW, SCHIFF MH, BAUMGARTNER SW *et al.*: Etanercept therapy in rheumatoid arthritis. A randomized, controlled trial. *Ann Intern Med* 1999; 130: 478-86.
50. VAN DER HEIJDE DM, VAN 'T HM, VAN RIEL PL, VAN DE PUTTE LB: Validity of single variables and indices to measure disease activity in rheumatoid arthritis. *J Rheumatol* 1993; 20: 538-41.
51. VAN DER HEIJDE DM, VAN 'T HM, VAN RIEL PL, VAN DE PUTTE LB: Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol* 1993; 20: 579-81.
52. PINCUS T, SOKKA T: Complexities in the quantitative assessment of patients with rheumatic diseases in clinical trials and clinical care. *Clin Exp Rheumatol* 2005; 23 (Suppl. 39): S1-S9.