LUND UNIVERSITY

Response criteria for rheumatoid arthritis in clinical practice - how useful are they?

Gülfe, Anders; Geborek, Pierre; Saxne, Tore

Link to publication

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Response criteria for rheumatoid arthritis in clinical practice: how useful are they?

A Gülfe, P Geborek and T Saxne

Updated information and services can be found at:
http://ard.bmjjournals.com/cgi/content/full/64/8/1186

These include:

References	This article cites 16 articles, 3 of which can be accessed free at: **http://ard.bmjjournals.com/cgi/content/full/64/8/1186#BIBL** 1 online articles that cite this article can be accessed at: **http://ard.bmjjournals.com/cgi/content/full/64/8/1186#otherarticles**
Rapid responses	You can respond to this article at: **http://ard.bmjjournals.com/cgi/eletter-submit/64/8/1186**
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article
Topic collections	Articles on similar topics can be found in the following collections Rheumatoid Arthritis (902 articles)

Notes

To order reprints of this article go to:
http://www.bmjjournals.com/cgi/reprintform

To subscribe to *Annals of the Rheumatic Diseases* go to:
http://www.bmjjournals.com/subscriptions/

EXTENDED REPORT

Response criteria for rheumatoid arthritis in clinical practice: how useful are they?

A Gülfe, P Geborek, T Saxne

Objective: To compare the performance of the American College of Rheumatology (ACR), European League Against Rheumatism (EULAR), and simple disease activity index (SDAI) response criteria for rheumatoid arthritis at the individual level in an observational cohort.

Methods: 184 outpatients were followed using a structured protocol. For each patient, the responses according to ACR 20% and 50%, EULAR moderate and good, and SDAI minor and major responses were calculated. For comparison, improvements in health assessment questionnaire (HAQ) score of 0.22 and 0.5 were calculated. The numbers of individuals fulfilling the criteria at each level were compared, and the numbers fulfilling any two sets of response criteria calculated. The EULAR "moderate" and "good" responses were grouped together as "overall," and SDAI "minor" and "major" were merged into SDAI "overall".

Results: All 94 ACR 20 responders were found in the EULAR and SDAI "overall" response groups, and 118 of 124 SDAI "overall" responders were found in the EULAR "overall" group. In contrast, of 53 ACR 50 responders, only 34 were found in the EULAR "good" or SDAI "major" group. Among the 56 patients in the EULAR "good" response group, only 26 met the SDAI "major" response. Improvement in HAQ score performed similarly to the other response criteria sets at the group levels.

Conclusions: For individual patients, agreement is good at the level of ACR 20 response, when EULAR overall, SDAI overall, or HAQ 0.22 criteria are applied. Agreement between ACR 50, EULAR good, SDAI major, and HAQ 0.5 response is poor. This should be considered when response criteria are used for clinical decisions.

See end of article for authors' affiliations

Correspondence to:
Dr Anders Gülfe,
Department of
Rheumatology, Lund
University Hospital, SE-
221 85 Lund, Sweden;
andersgulfe@hotmail.com

Accepted 30 January 2005
**Published Online first
10 March 2005**

Rheumatoid arthritis is a chronic, disabling disease affecting about 0.5% of the population.[1] Pharmacological treatment tends to be of long duration and may be complex. Response is often suboptimal, and toxic side effects are not uncommon. No single measure of disease activity or changes in activity (that is, the difference in disease activity between two observations) has proven sufficient, and a variety of composite indices have thus been developed. The utility of such standardised response criteria, for example the ACR (American College of Rheumatology) 20–50–70% response[2] and the European League Against Rheumatism (EULAR) response criteria[3] are well established for use in clinical trials, where the proportion of patients responding constitutes a measure of efficacy compared to placebo or a standard treatment such as methotrexate. This practice has greatly facilitated the evaluation of novel treatments. The disease activity score (DAS) and its variants,[4 5] and the simple disease activity index (SDAI),[6] are intended for routine clinical use. The components of the various response criteria sets are shown in table 1.

Various response criteria sets have been validated against each other in randomised, controlled trials (RCTs) of antirheumatic treatment regimens.[7] In general, the degree of agreement between different response criteria sets is fair in RCTs. The problem of different responsiveness at the individual patient level is, however, seldom addressed.

In a previous communication we described a clinical protocol for monitoring treatment in patients with rheumatoid arthritis.[8] The protocol is suitable for monitoring patients seen in routine clinical practice and can be used to estimate the efficacy and tolerability of different treatment regimens in spite of possible confounding by indication. The individual patient's reaction to treatment according to sets of response criteria is easily determined using this protocol. We have recently reported the ACR responses for etanercept, inflximab, and leflunomide in patients completing the first year of treatment.[9]

Our aim in the present study was to compare the performance of the ACR, EULAR, and SDAI response criteria sets in individual patients, in an observational study of patients with long standing established rheumatoid arthritis treated with tumour necrosis factor α (TNFα) blockers. Improvement in a patient administered instrument, the health assessment questionnaire (HAQ), was used for comparison.

METHODS

Patients attending the Department of Rheumatology, Lund University Hospital, and who had started treatment with TNFα blockers were entered consecutively into a database.

Requirements for inclusion in the study were a diagnosis of rheumatoid arthritis according to the ACR 1987 diagnostic criteria,[10] and treatment failure on at least two disease modifying antirheumatic drugs (DMARDs) including methotrexate. The patients had to be included in the database between 1 January 1999 and 31 December 2001, and a complete dataset at baseline and at three months had to be available. To investigate whether the response pattern changed with longer treatment time, similar analyses were also carried out for patients with a complete dataset at six months.

The protocol comprises the following variables: the 28 joint swollen and tender joint count, patient's global visual

Abbreviations: ACR, American College of Rheumatology; DAS, disease activity score; EULAR, European League Against Rheumatism; HAQ, health assessment questionnaire; SDAI, simple disease activity index

Table 1 Components of the various response criteria sets

Criteria set	Tender joint count	Swollen joint count	Patient global VAS	Patient pain VAS	Evaluator's global	HAQ	ESR	CRP
ACR	+	+	+/−	+/−	+/−	+/−	+/−	+/−
EULAR (DAS)	+	+	+	−	−	−	+	−
SDAI	+	+	+	−	+	−	−	+

+, required; −, not required; in the ACR response criteria, any three of the variables marked "+/−" are required. For details about the response criteria, see references 1–5.
ACR, American College of Rheumatology; CRP, C reactive protein; DAS, disease activity score; ESR, erythrocyte sedimentation rate; EULAR, European League Against Rheumatism; SDAI, simple disease activity index; VAS, visual analogue scale.

analogue scale (VAS) (a 10 cm non-anchored horizontal line[11]), patient's pain VAS, the health assessment questionnaire (HAQ),[12][13] and the evaluator's global assessment of disease activity (five degrees: inactive, low, moderate, high, or maximal), erythrocyte sedimentation rate (ESR) according to Westergren, and C reactive protein.

These variables were used to calculate fulfilment of the following response criteria: ACR 20% and 50%; EULAR non-response, moderate response, and good response; SDAI minor and major responses; and improvement in HAQ score of 0.22 (HAQ 0.22) and 0.5 (HAQ 0.5). The reason for using these HAQ levels of improvement are that 0.22 has been shown to be a level of improvement perceived beneficial by the patient,[14] and 0.5 has been used in health economics models.[15] For the purpose of the present study, EULAR moderate and good responders and SDAI minor and major were grouped together as "overall" responders. The numbers of individuals fulfilling the respective response criteria at each level were compared and the agreement between individual patients fulfilling two sets of response criteria was calculated for each possible pair of response criteria at the actual level.

RESULTS

During the period 184 rheumatoid patients fulfilled the requirements for evaluation in the study. The characteristics of the patients at baseline and for some of the variables at three and six months are shown in table 2.

For ACR 20, EULAR overall, SDAI overall, and HAQ 0.22 response, the proportion of responders was 51%, 74%, 67%, and 64%, respectively. All 94 ACR 20 patients were found in the EULAR overall and the SDAI overall groups, while the
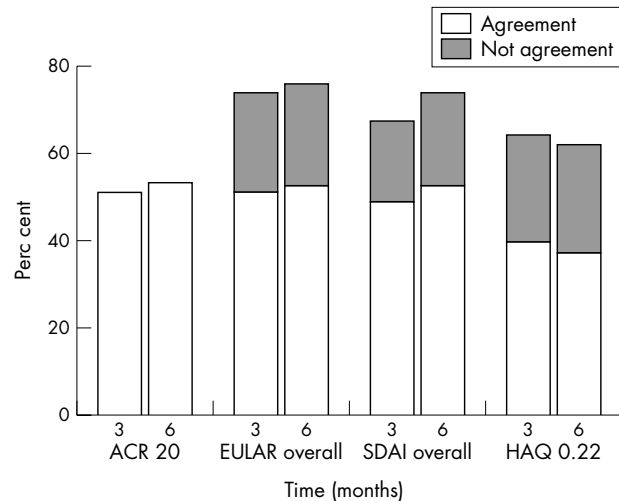
Figure 1 Proportion of patients fulfilling less strict response criteria sets at three and six months. The agreement between the different criteria sets using ACR 20% as reference is demonstrated.

HAQ 0.22 showed agreement in 73 of these patients. The absolute majority of the SDAI overall (118/124) are found in the EULAR overall group, which comprises 136 patients. For HAQ improvement 0.22 the agreement with EULAR overall is 94/118 and for SDAI overall 90/118 (table 3, fig 1).

For ACR 50, EULAR good, SDAI major, and HAQ 0.5 response, the response rates were 29%, 30%, 30%, and 29%, respectively. However, at the individual level only 34 of the 53

Table 2 Characteristics of the cohort (n = 184) at inclusion and at three and six months

	Baseline	3 Months	6 Months
Numbers	184	184	150
Male/female	46/138		
Disease duration (years)	12 (0 to 55)		
Age at inclusion (years)	56 (20 to 84)		
Previous number of DMARDs	3 (2 to 5)		
Steroid dose (mg/week)	35 (17 to 57)	35 (0 to 53)*	18 (0 to 35)*
ESR (mm/h)	31.5 (20 to 54)	20 (12 to 34)*	20 (10 to 38)*
C reactive protein (mg/litre)	21 (10 to 42)	9 (1 to 22)*	9 (1 to 23)*
VAS pain (mm)	64 (47 to 78)	34 (16 to 55)*	29 (15 to 55)*
VAS global (mm)	66.5 (49 to 81)	34 (15 to 55)*	29 (15 to 57)*
Physician's global assessment	2 (2 to 3)	1 (1 to 2)*	1 (1 to 2)*
28 tender joint count	7.5 (3 to 13)	2 (0 to 5.5)*	2 (0 to 6)*
28 swollen joint count	9 (5 to 12)	3 (1 to 6)*	3 (0 to 6)*
HAQ	1.4 (1 to 1.9)	1.1 (0.6 to 1.6)*	1.1 (0.6 to 1.6)*
DAS 28	5.6 (4.7 to 6.4)	3.8 (2.8 to 4.8)*	3.7 (2.7 to 4.8)*
SDAI	31 (24 to 41)	14 (9 to 24)*	13 (7 to 22)*

Values for median and (25th to 75th centile). Physician's global assessment is recorded at a 5 point Likert scale.
*p<0.001 v baseline.
DAS 28; 28 joint disease activity score; DMARD, disease modifying antirheumatic drug; ESR, erythrocyte sedimentation rate; HAQ, health assessment questionnaire; RF, rheumatoid factor; SDAI; simple disease activity index; VAS, visual analogue scale.

Table 3 Agreement of response criteria fulfilment in individual patients

	Total responders	Agreement (n)						
		ACR 20	EULAR overall	SDAI overall	HAQ 0.22	ACR 50	EULAR good	SDAI major
ACR 20	94							
EULAR overall	136	94						
SDAI overall	124	90	118					
HAQ 0.22	118	73	94	87				
ACR 50	53	53	53	51	44			
EULAR good	56	46	56	50	42	34		
SDAI major	56	49	56	56	38	34	26	
HAQ 0.5	55	30	48	47	55	31	28	31

For details, see text.
ACR 20, American College of Rheumatology criteria, 20% response; EULAR, European League Against Rheumatism; HAQ 0.5, improvement by 0.5 in health assessment questionnaire; SDAI; simple disease activity index.

ACR 50 responders fulfilled EULAR good or SDAI major responses or both. The EULAR good response agreement with SDAI major response was only 26/56. HAQ 0.5 agreement was of the same magnitude (28/53 *v* EULAR good response, 34/53 *v* SDAI major response (table 3, fig 2)).

At six months, data from 150 patients were available (table 2). Agreement was similar (examples shown in figs 1 and 2).

DISCUSSION

Evaluation of treatment effects in rheumatoid arthritis has received more attention with the introduction of novel therapies, notably the TNFα blockers, which are very expensive and for which the long time effects are unknown. Standardised measures of efficacy should be reliable and simple to use in everyday practice. The criteria sets studied contain much the same variables, but they are weighted or handled somewhat differently. The history and philosophy behind the criteria sets are also different. It is therefore interesting that the agreement between the three criteria sets at the individual patient level was fair at the lower levels of response. More patients respond on the EULAR and SDAI scales than on ACR 20, but the agreement was very good, indicating that, in this observational cohort, patients responded similarly on all three criteria sets.

HAQ, a measure of function, and HAQ improvement have been employed for testing construct validity of the EULAR and SDAI criteria, as it is not included in these.[3][6] It contributes only little to the ACR response criteria set. The

Figure 2 The proportion of patients fulfilling more strict response criteria sets at three and six months. Agreement using EULAR good response is given as reference.

fixed levels of 0.22 and 0.5 were in some aspects arbitrarily chosen, but they have been used in previous studies. One reason for using HAQ improvement is the concept that patient self report questionnaires are sufficient to evaluate the efficacy of rheumatoid arthritis treatment.[16] Indeed, at the levels chosen, HAQ improvement did not perform very differently from the other criteria sets tested at the group level in our study.

Results were different when comparing the ACR 50 with the EULAR good and the SDAI major responses, respectively. At the group level, they performed similarly—that is, the same numbers of patients tended to respond, irrespective of the criteria sets applied. When individual responses were analysed, however, agreement was poor. At the higher level of response, EULAR good and SDAI major showed agreement in only 34 of 56 patients. If used as a basis for treatment decisions in the individual patient, the choice of criteria would have a major impact. The clinical importance of this may be limited, given the fact that many patients do not respond at this level in routine care, and that the lower degree of response will often be considered sufficient to continue with a particular treatment. However, as treatment modes become more effective, goals of therapy will change, and aiming at remission or near remission will be increasingly realistic. Thus the verification of response at higher levels will become more important.

Standardised response criteria and activity scores in rheumatoid arthritis have proven their utility in clinical trials, when groups of patients are analysed statistically and biological variation tends to be levelled out. They also perform well in observational studies to estimate the response at group level, and in this context they can be indicative of the efficacy of various treatments in clinical practice. In this non-randomised, observational cohort of long standing rheumatoid patients treated with TNFα blockers, the various criteria sets appeared to perform differently at the individual level at the higher degree of response. It may thus be wise to consider response criteria fulfilment in individuals with some scepticism and not to rely heavily on them in clinical practice, but to look upon them as one of several aids for treatment decisions. Absolute measures of disease activity, such as DAS or SDAI levels, are probably better for treatment decisions in the daily care of individual patients. Clinical judgement remains crucial in the management of rheumatoid patients, but currently available response criteria, although not perfect, may be included in the evaluation of treatment with antirheumatic drugs to facilitate monitoring of treatment response.
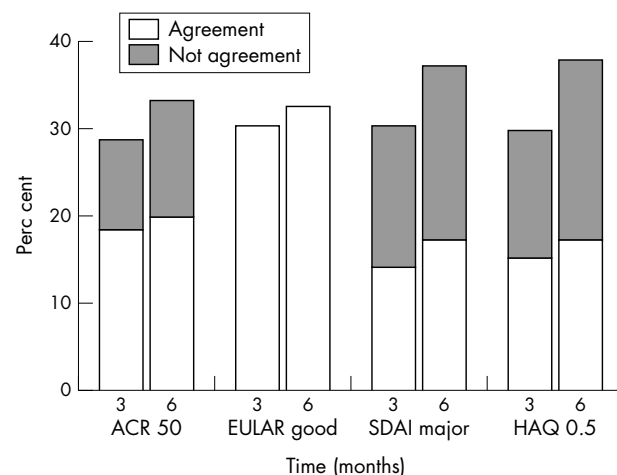
....
Authors' affiliations
A Gülfe, P Geborek, T Saxne, Department of Rheumatology, Lund University Hospital, Lund, Sweden

REFERENCES

1 **Simonsson M**, Bergman S, Jacobsson LT, Petersson IF, Svensson B. The prevalence of rheumatoid arthritis in Sweden. *Scand J Rheumatol* 1999;**28**:340–3.

2 **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.

3 **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.

4 **van der Heijde DM**, van't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, *et al.* Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;**49**:916–20.

5 **Prevoo 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.

6 **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.

7 **Verhoeven AC**, Boers M, van Der LS. Responsiveness of the core set, response criteria, and utilities in early rheumatoid arthritis. *Ann Rheum Dis* 2000;**59**:966–74.

8 **Geborek P**, Saxne T. Clinical protocol for monitoring of targeted therapies in rheumatoid arthritis. *Rheumatology (Oxford)* 2000;**39**:1159–61.

9 **Geborek P**, Crnkic M, Petersson IF, Saxne T, South Swedish Arthritis Treatment Group. Etanercept, infliximab, and leflunomide in established rheumatoid arthritis: clinical experience using a structured follow up programme in southern Sweden. *Ann Rheum Dis* 2002;**61**:793–8.

10 **Arnett FC**, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;**31**:315–24.

11 **Ferraz MB**, Quaresma MR, Aquino LR, Atra E, Tugwell P, Goldsmith CH. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. *J Rheumatol* 1990;**17**:1022–4.

12 **Ekdahl C**, Eberhardt K, Andersson SI, Svensson B. Assessing disability in patients with rheumatoid arthritis. Use of a Swedish version of the Stanford Health Assessment Questionnaire. *Scand J Rheumatol* 1988;**17**:263–71.

13 **Fries JF**, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;**23**:137–45.

14 **Wells GA**, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol* 1993;**20**:557–60.

15 **Kobelt G**, Jonsson L, Young A, Eberhardt K. The cost-effectiveness of infliximab (Remicade) in the treatment of rheumatoid arthritis in Sweden and the United Kingdom based on the ATTRACT study. *Rheumatology (Oxford)* 2003;**42**:326–35.

16 **Pincus T**, Strand V, Koch G, Amara I, Crawford B, Wolfe F, *et al.* An index of the three core data set patient questionnaire measures distinguishes efficacy of active treatment from that of placebo as effectively as the American College of Rheumatology 20% response criteria (ACR20) or the Disease Activity Score (DAS) in a rheumatoid arthritis clinical trial. *Arthritis Rheum* 2003;**48**:625–30.