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

Extended report

Anders Gülfe, MD, Pierre Geborek, MD, PhD, Tore Saxne, MD, PhD

Department of Rheumatology, Lund University Hospital, SE-221 85 Lund, Sweden

Corresponding author: Anders Gülfe, Department of Rheumatology, Lund University Hospital, SE-221 85 Lund, Sweden, tel: +46 46171619, fax: +46 46128468, e-mail: andersgulf@hotmail.com
Abstract

Objective. Comparison of the performance of the ACR, EULAR and SDAI response criteria for RA at the individual level, in an observational cohort.

Methods. We followed 184 outpatients according to a structured protocol. For each patient, the responses according to ACR20 and 50%, EULAR moderate and good, and SDAI minor and major responses, were calculated. For comparison also improvement in HAQ score of 0.22 and 0.5 were calculated. The numbers of individuals fulfilling the criteria at each level were compared, and the number of individuals fulfilling any two sets of response criteria was 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 118/124 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.

Conclusion. 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.

Key words
Rheumatoid arthritis; treatment; response criteria; individual response; ACR20-50% response; EULAR response criteria; Simple Disease Activity Score.
Introduction
Rheumatoid arthritis (RA) is a chronic, disabling disease affecting about 0.5% of the population[1]. Pharmacological treatment tends to be of long duration and is often complex. Response is often suboptimal, and toxic side effects are not uncommon. No single measure of disease activity or changes thereof (i.e. 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 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, e.g. 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.

Table 1. Components of the various response criteria sets. + = 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.

<table>
<thead>
<tr>
<th>Criteria set</th>
<th>Tender joint count</th>
<th>Swollen joint count</th>
<th>Patient global VAS</th>
<th>Patient pain VAS</th>
<th>Evaluator’s global HAQ</th>
<th>ESR</th>
<th>CRP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACR</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>EULAR (DAS)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>SDAI</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Various response criteria sets have been validated against each other in randomised, controlled trials (RCTs) of anti-rheumatic 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 have described a clinical protocol for monitoring treatment of patients with RA[8]. The protocol is suitable for monitoring patients seen in routine clinical practice, and it can be used to estimate the efficacy and tolerability of different treatment regimens in spite of a possible confounding by indication. The individual patient’s reaction to treatment according to sets of response criteria is easily determined by means of the protocol. We have recently reported the ACR-responses for etanercept, infliximab and leflunomide for patients completing the first year of treatment[9].

The objective of the present study is to compare the performance in individual patients of the ACR, EULAR and SDAI response criteria sets, in an observational study of patients with long-standing, established RA, treated with TNFα-blockers. Also, improvement in a patient administered instrument, HAQ, was used for comparison.

Materials and methods
Patients attending the Department of Rheumatology, Lund University Hospital, who started treatment with TNFα-blockers were entered consecutively into a database.
Requirements for inclusion in the present study were a diagnosis of RA according to the ACR 1987 diagnostic criteria[10], and treatment failure on at least two disease modifying anti-rheumatic drugs (DMARDs) including methotrexate. Furthermore, the patients had to be included in the database between January 1, 1999 and December 31, 2001 and a complete data set at baseline and at 3 months should be available. To investigate if the response pattern changed with longer treatment time, similar analyses were performed also for patients with a complete data set at 6 months.

The protocol comprises the following variables: 28-joint swollen and tender joint count, patient’s global VAS, patient’s pain VAS (VAS = visual analogue scale, 10 cm non-anchored horizontal line[11], Health Assessment Questionnaire (HAQ)[12;13] and evaluator’s global assessment of disease activity (5 degrees - non, low, moderate, high or maximal), ESR (erythrocyte sedimentation rate) according to Westergren and CRP (C-reactive protein). These variables are used to calculate fulfilment of the following response criteria: ACR20 and 50%, EULAR non-, moderate- and good, 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 proved to be a level of improvement perceived beneficial by the patient [14], and 0.5 has been used in health economic 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 of 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 RA patients fulfilled the requirements for evaluation in the study. The characteristics of the patients at baseline and for some of the variables at 3 and 6 months are shown in Table 2.

Table 2. Characteristics of the cohort (N=184) at inclusion and 3 and 6 months. Values for median and (25th-75th percentile) are given. RF = Rheumatoid factor. Physician’s global assessment is recorded at a 5 point Likert scale. *** = P<0.001 compared to baseline

<table>
<thead>
<tr>
<th>Number</th>
<th>Baseline</th>
<th>3 months</th>
<th>6 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male/female</td>
<td>184</td>
<td>184</td>
<td>150</td>
</tr>
<tr>
<td>Disease duration (yrs)</td>
<td>12 (0-55)</td>
<td>12 (0-55)</td>
<td>12 (0-55)</td>
</tr>
<tr>
<td>Age at inclusion</td>
<td>56 (20-84)</td>
<td>56 (20-84)</td>
<td>56 (20-84)</td>
</tr>
<tr>
<td>Previous no of DMARDs</td>
<td>3 (2-5)</td>
<td>3 (2-5)</td>
<td>3 (2-5)</td>
</tr>
<tr>
<td>Steroid dose (mg/week)</td>
<td>35 (17-57)</td>
<td>35 (17-57)</td>
<td>35 (17-57)</td>
</tr>
<tr>
<td>ESR (mm in 1st hour)</td>
<td>31.5 (20-54)</td>
<td>31.5 (20-54)</td>
<td>31.5 (20-54)</td>
</tr>
<tr>
<td>C-reactive protein (mg/litre)</td>
<td>21 (10-42)</td>
<td>21 (10-42)</td>
<td>21 (10-42)</td>
</tr>
<tr>
<td>VAS pain (mm)</td>
<td>64 (47-78)</td>
<td>64 (47-78)</td>
<td>64 (47-78)</td>
</tr>
<tr>
<td>VAS global (mm)</td>
<td>66.5 (49-81)</td>
<td>66.5 (49-81)</td>
<td>66.5 (49-81)</td>
</tr>
<tr>
<td>Physician’s global assessment</td>
<td>2 (2-3)</td>
<td>2 (2-3)</td>
<td>2 (2-3)</td>
</tr>
<tr>
<td>28 tender joint count</td>
<td>7.5 (3-13)</td>
<td>7.5 (3-13)</td>
<td>7.5 (3-13)</td>
</tr>
<tr>
<td>28 swollen joint count</td>
<td>9 (5-12)</td>
<td>9 (5-12)</td>
<td>9 (5-12)</td>
</tr>
<tr>
<td>HAQ</td>
<td>1.4 (1-1.9)</td>
<td>1.4 (1-1.9)</td>
<td>1.4 (1-1.9)</td>
</tr>
<tr>
<td>DAS28</td>
<td>5.6 (4.7-6.4)</td>
<td>5.6 (4.7-6.4)</td>
<td>5.6 (4.7-6.4)</td>
</tr>
<tr>
<td>SDAI</td>
<td>31 (24-41)</td>
<td>31 (24-41)</td>
<td>31 (24-41)</td>
</tr>
</tbody>
</table>
For ACR 20, EULAR overall and SDAI overall and HAQ 0.22 response, the proportion of responders is 51%, 74%, 67%, and 64% respectively. All 94 ACR 20 patients are found in the EULAR overall, and SDAI overall groups, while the HAQ 0.22 shows 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 are 29%, 30%, 30%, and 29%, respectively. However, at the individual level only 34/53 of the ACR 50 responders fulfill EULAR good and/or SDAI major response. The EULAR good response agreement with SDAI major response is only 26/56. HAQ 0.5 agreement is of the same magnitude with 28/53 versus EULAR good response, and 34/53 versus SDAI major response (Table 3, Fig 2).

**Table 3.** Agreement of response criteria fulfilment in individual patients. For details, see text.

<table>
<thead>
<tr>
<th>Total responders</th>
<th>Agreement (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N ACR20</td>
</tr>
<tr>
<td>ACR20</td>
<td>94</td>
</tr>
<tr>
<td>EULAR overall</td>
<td>136</td>
</tr>
<tr>
<td>SDAI overall</td>
<td>124</td>
</tr>
<tr>
<td>HAQ 0.22</td>
<td>118</td>
</tr>
<tr>
<td>ACR50</td>
<td>53</td>
</tr>
<tr>
<td>EULAR good</td>
<td>56</td>
</tr>
<tr>
<td>SDAI major</td>
<td>56</td>
</tr>
<tr>
<td>HAQ 0.5</td>
<td>55</td>
</tr>
</tbody>
</table>

At 6 months, data from 150 patients were available (table 2). Agreement was similar (examples shown in fig 1 and 2).

**Discussion**

Evaluation of treatment effects in RA has received more attention with the introduction of novel therapies, notably the TNFα blockers, which are very expensive and whose 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 is fair, at the lower levels of response. More patients respond according to the EULAR and SDAI, than to the ACR 20, but the agreement is very good, indicating that, in this observational cohort, similar patients respond according to all three criteria sets.

HAQ, a measure of function, and HAQ improvement has 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 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 efficacy of RA treatment[16]. Indeed, at the levels chosen, HAQ improvement did not perform much differently from the other criteria sets tested at the group level in our study.

Results are different when comparing the ACR 50 to the EULAR good and the SDAI major responses, respectively. At the group level, they perform similarly, i.e. the same number of patients tend to respond, irrespective of the criteria sets applied. When individual responses are analysed, however, agreement is poor. At the higher level of response, EULAR good and SDAI major show agreement in only 34/56 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 often will be considered sufficient to continue therapy. However, as treatment modalities become more effective, goals of therapy will change, and aiming at remission or near-remission will be increasingly realistic. Thus verification of response at higher levels will become more important.

Standardised response criteria and activity scores in RA 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 RA patients treated with TNFα blockers, the various criteria sets appear to perform differently on 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 RA 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.

Acknowledgments
This study was supported by grants from Österlund and Kock Foundations, King Gustav V 80 year fund, and Reumatikerförbundet.

This study was approved by the Ethics Committee, Lund University.

Competing interests: None.

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd and its Licensees to permit this article (if accepted) to be published in ARD and any other BMJPGL products to exploit all subsidiary rights, as set out in our licence (http://ard.bmjournals.com/misc/ifora/licenceform.shtml).
Legends

**Figure 1.** Proportion of patients fulfilling less strict response criteria sets at 3 and 6 months. The agreement between the different criteria sets using ACR20% as reference is demonstrated.

**Figure 2.** Proportion of patients fulfilling more strict response criteria sets at 3 and 6 months. Agreement using EULAR good response as reference.
Reference List


Figure 2

The chart shows the percentage of agreement and not agreement over time for different criteria:

- **ACR 50**
  - 3 months: 20%, 6 months: 30%

- **EULAR good**
  - 3 months: 25%, 6 months: 30%

- **SDAI major**
  - 3 months: 35%, 6 months: 40%

- **HAQ 0.5**
  - 3 months: 30%, 6 months: 35%

The x-axis represents the months (3 and 6), and the y-axis represents the percentage range from 0 to 40.
Response criteria for rheumatoid arthritis in clinical practice - how useful are they?

Anders Gülfe, Pierre Geborek and Tore Saxne

Ann Rheum Dis published online March 10, 2005

Updated information and services can be found at:
http://ard.bmj.com/content/early/2005/03/10/ard.2004.027649.citation

Email alerting service

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

Topic Collections

Articles on similar topics can be found in the following collections

- Connective tissue disease (4253)
- Degenerative joint disease (4641)
- Immunology (including allergy) (5144)
- Musculoskeletal syndromes (4951)
- Rheumatoid arthritis (3258)

Notes

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

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

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