Defining response to disease modifying antirheumatic drugs in patients with rheumatoid arthritis.

Samantha L Hider, Ian N Bruce, Alan J Silman and Deborah P M Symmons

J Rheumatol 2005;32;6-10
http://www.jrheum.org/content/32/1/6.citation

1. Sign up for TOCs and other alerts
   http://www.jrheum.org/alerts

2. Information on Subscriptions
   http://jrheum.com/faq

3. Information on permissions/orders of reprints
   http://jrheum.com/reprints_permissions

The Journal of Rheumatology is a monthly international serial edited by Earl D. Silverman featuring research articles on clinical subjects from scientists working in rheumatology and related fields.
Defining Response to Disease Modifying Antirheumatic Drugs in Patients with Rheumatoid Arthritis

In recent years there have been considerable advances in the pharmacological management of rheumatoid arthritis (RA). Previously the mainstay of treatment for RA was disease modifying antirheumatic drugs (DMARD) that suppress the inflammatory process and slow disease progression. Today’s biologic agents are specifically designed to block key mediators of the RA inflammatory process. The efficacy of these agents, both in terms of disease activity and radiological progression, is now well established. However these drugs are expensive and, where resources are limited, it is important to target their use in those patients most likely to benefit. In the United Kingdom, for example, the National Institute for Clinical Excellence (NICE) Guidelines currently recommend that biologic agents should only be considered in patients with active RA who have failed to respond to at least 2 DMARD (including methotrexate, MTX). The NICE recommendation has focused attention on how treatment response should be defined in RA, and whether it is possible to predict which patients will respond to a particular DMARD. In this article we consider the ways in which DMARD treatment response can be defined, and in an accompanying article we consider the known predictors of DMARD response.

RESPONSE AS AN INDIVIDUAL OR GROUP CHARACTERISTIC

It is well recognized that response to DMARD is variable; it has been estimated that, in clinical practice, only around 50% of patients respond adequately to any one traditional DMARD even in the short term. Moreover, a substantial number of patients have to stop DMARD because of side effects before they have had a chance to respond. It is likely that clinical response to DMARD is a continuous variable, but until recently rheumatology clinical trials have tended to compare the average response. By contrast, in clinical practice it is necessary to focus on response at the individual level and to define response/non-response using an arbitrary threshold across one or more outcome measures. Based on such an approach 3 groups are recognized: non-responders (in whom the drug is stopped), partial responders (in whom a second DMARD or steroids are added), and responders (in whom the original DMARD is continued). Recent clinical trials have begun to classify patients as responders or non-responders to drugs in a similar way. However there are more limited data on “partial responders.” Although combining DMARD after partial response to the first agent is a common strategy in clinical practice, few trials have considered such an approach, except for some biologics trials when these agents were prescribed to patients with partial response to MTX. Therefore, for both clinical trials and routine practice, it is necessary to develop meaningful cutoff points between response, partial response, and non-response.

CURRENT DEFINITIONS OF DMARD RESPONSE

Most attempts to classify DMARD response have been set in the context of clinical trials. A number of approaches have been tried. The first (the American Rheumatism Association RA remission criteria) were based on defining complete remission. One hundred and seventy-five patients considered to be in complete remission (including 112 receiving treatment) were compared with 169 patients either in partial remission or with active disease. Combinations of 6 individual criteria yielded optimal discrimination between the 2 groups (Table 1). In the study sample, 5 or more of these criteria in an individual patient yielded 72% sensitivity and 100% specificity for complete remission. While the end-point of complete remission is obviously desirable, it is clear that, even with the new biologic agents, complete remission of RA is seldom achieved. As a consequence, it has proved necessary to develop criteria to identify those patients who have experienced a worthwhile improvement on therapy.

The second approach was based on clinical trial data; statistical techniques were used to identify the level of response of an effective DMARD needed to be superior to placebo. Three criteria sets have been proposed, 2 of which are based on a relative change rather than achievement of a particular disease state. The first, published by Paulus, et al, was derived from an arbitrarily selected group of variables. The improvement criteria were thus based on which of these many variables had to improve, and to what extent, in order to discriminate between active drug and placebo in...
multicenter DMARD clinical trials. As shown in Table 2, most variables had to improve by at least 20% in the active treatment arm. For the second criteria set, the American College of Rheumatology (ACR) committee on Outcome Measures in RA Clinical Trials used a similar approach (Table 2)\(^1\). They compiled a dataset of 69 “paper patients” with data at baseline and at 6 months “post treatment” on all items included in the core set of outcome measures recommended for use in RA clinical trials\(^1\). This was sent to 89 rheumatologists who were asked to state whether they thought that the patient had improved following treatment with the DMARD. From the 68 rheumatologists’ responses a number of possible definitions of improvement were derived. These were then validated using datasets from 5 placebo-controlled DMARD clinical trials. The definition of improvement chosen was that which categorized the fewest placebo treated patients as “improved.” However, the resulting definition, (which has become known as the ACR20) identifies patients who have shown any response to a DMARD rather than those who have shown an adequate response.

Most patients with RA and their physicians are, however, looking for an improvement of more than 20% when starting a DMARD. Recent trials have therefore reported on the number of patients who have achieved 50% (ACR50) or 70% (ACR70) improvement. The ACR50 requires a patient to have improved by at least 50% in both swollen and tender joint counts and by at least 50% in 3 of the remaining 5 core set measurements (Table 2). The ACR70 has a similar definition based on 70% improvement\(^2\). Felson, et al\(^2\) found, when analyzing clinical trial data, that using the ACR50 or ACR70 as the definition of response led to a greater fall in response rates in the active treatment arms than in the placebo arms. Thus these definitions were less useful than the ACR20 in identifying patients with any response to treatment. The clinical utility of this approach remains to be established\(^3\). Thus although the ACR20 is a better index to discriminate between active agent and placebo, for current drug trials it may be more relevant to consider the ACR50 or ACR70.

The ACR response criteria measure change from baseline rather than the state at the end of the trial. At the recent OMERACT 7 (Outcome Measures in Rheumatology Clinical Trials) meeting, an attempt was made to define a “low disease activity state” (LDAS) using the ACR/OMERACT core set of RA outcome measures. LDAS was defined as “that disease activity state which is deemed acceptable by the physician and patient, given currently available treatment and its limitations.” Specific values of the core set variables were derived by consensus.

The European League Against Rheumatism (EULAR) derived response criteria using the Disease Activity Score (DAS)\(^4\). This is a composite measure of disease activity modelled on the physician’s decision to start DMARD therapy in patients with early RA\(^5\). The original DAS was a complex formula that incorporated the Ritchie articular index (based on 53 joints), a swollen joint count (based on 44 joints), erythrocyte sedimentation rate (ESR), and the patient global assessment.

Subsequent modification showed that an equivalent formula could be developed using the 28-joint tender and swollen joint counts in place of the Ritchie articular index and the 44-joint swollen joint count, respectively (the so-called DAS28 score)\(^6\), which has since been validated\(^7\). The DAS and DAS28 are not directly interchangeable as they have different ranges. A change of 1.2 points in either

### Table 1.  
American Rheumatism Association (ARA) remission criteria (1981)\(^8\). Five or more criteria must be present for at least 2 consecutive months.

1. Morning stiffness less than 15 minutes
2. No fatigue
3. No joint pain (by history)
4. No joint tenderness or pain on movement (by examination)
5. No soft tissue swelling in joints or tendon sheaths
6. ESR < 50 mm/h (females) or < 20 mm/h (males)

ESR: erythrocyte sedimentation rate.

### Table 2.  
Response criteria based on statistical tests to distinguish an active DMARD from placebo.

<table>
<thead>
<tr>
<th></th>
<th>Paulus Criteria(^9)</th>
<th>ACR20 Criteria(^1)</th>
<th>ACR50 Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement (%) From Baseline in:</td>
<td>4 out of 6</td>
<td>First 2 Criteria plus 3 of Remaining 5</td>
<td>First 2 Criteria plus 3 of Remaining 5</td>
</tr>
<tr>
<td>Tender joint count</td>
<td>20</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Swollen joint count</td>
<td>20</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>ESR</td>
<td>20</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Physician global DAS</td>
<td>40</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Patient global DAS</td>
<td>40</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Duration of morning stiffness</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain score</td>
<td>20</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Physical function</td>
<td>20</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

DAS: Disease Activity score.
DAS or DAS28 (2 times the measurement error) is considered statistically significant. Patients are then classified as good, moderate, or non-responders (Table 3) according to both a significant change in the DAS and the level of residual disease activity.

Thus the “change” component of the EULAR response criteria was chosen to detect a statistically significant change rather than an adequate response. The criteria have the added face validity of incorporating knowledge of the current activity “state.” However, the definition of low disease activity state (DAS ≤ 2.4 or DAS28 ≤ 3.2) is not equivalent to disease remission. By contrast, a DAS score of ≤ 1.6 (equivalent to a DAS28 score of ≤ 2.6) has been shown to be the equivalent of remission according to the ARA remission criteria in a cohort of patients with early RA21.

All 3 criteria sets therefore require a baseline assessment of disease activity, whether relative or absolute change is being considered. This may be appropriate when examining response over the short term (less than 12 mo), but not over a longer period. In addition, since all these measures are composite scores, patients may meet the threshold for improvement by different routes. Some patients may be considered partial responders if they reach the threshold for improvement in some but not all of the criteria required. This may be a particular issue in established disease where disease damage may mean that factors such as pain or physical function indices may be less sensitive to change. Also, although all 3 sets of criteria consider disease activity at 2 time points, they do not consider either the rate or duration of improvement.

To take account of duration of improvement in clinical trials, several studies have examined the utility of using area under the curve (AUC) measurements22,23. AUC converts serial measurements into a summary measure. However, to achieve this actual values of the clinical measures rather than cutoffs must be used. Attempts have been made to combine such an approach with the multidimensional outcome assessment of the ACR type. The numeric ACR (ACR-N) at a single point in time is defined as the lowest percentage improvement from baseline of 3 measures: tender joint count, swollen joint count, and median improvement of the remaining 5 core measures (Table 2) (i.e., if the 5 remaining core variables improved by 52%, 51%, 49%, 22%, and 47%, respectively, then the value used would be 49%)22,23. The AUC for the ACR-N can then be plotted. AUC can also be calculated for the DAS24. While such procedures are useful in clinical trials and in longitudinal observational studies25, their value in routine clinical practice is less evident. Also the AUC may be misleading, since depending on when the cutoff is taken, similar results may be obtained between a drug that works quickly but then becomes less effective and a drug that takes longer to work well. To date there are no accepted criteria for treatment response using the ACR-N.

In summary, the existing drug response criteria sets have been developed with the aim of either distinguishing an active DMARD from placebo or of detecting the minimum improvement that might be statistically significant. The EULAR response criteria based on the DAS (Table 3) have the added advantage of setting a threshold of disease activity below which a patient should fall in order to be judged a good responder. However, this choice of maximum level of residual disease activity was not based on the patient’s ultimate disease outcome, and this criteria set was developed in patients with early RA.

THE WAY FORWARD

In any clinical or trial context the definitions of treatment response and non-response have to be mirror images (Figure 1). Thus in Figure 1A the ACR20 are used to separate those who have failed to respond to a DMARD from those who have shown any response to the drug. The main use of the ACR20 in clinical practice is likely to be in defining primary non-response. In Figure 1B a definition is chosen that focuses on “satisfactory” or “adequate” response, which reflects the approach in clinical practice. In this context,
DMARD therapy is to retard disease progression. It may be terms of longer term outcomes since the ultimate aim of definitions at OMERACT 7.

To MTX would you continue this patient's current medication? If other DMARD were available of equivalent efficacy and toxicity providing the physician with a scenario (e.g., “if a new treatment”) would you choose to continue this patient's current medication? One approach to define “satisfactory” response would be to model the physician’s decision to stop or continue the DMARD (in the same way as the DAS was modelled on the decision to start treatment). However, such decisions are context and patient specific. Thus, in an era when there was only one DMARD, patients would continue the drug for much longer than when there were another 6 to 8 DMARD to try. Similarly the individual patient may continue their 5th for longer than their 1st DMARD, even if the improvement observed is less, because there are fewer remaining options for treatment. In addition, drugs such as MTX are now being used at higher doses, so a patient considered to be a non-responder to MTX at 10 mg orally per week might respond to a higher (or parenteral) dose. The decision to continue a DMARD is also influenced by whether the patient is experiencing any side effects. Thus 2 patients may show a similar moderate response to a DMARD — and that DMARD would be continued in the patient who felt well, and discontinued in the patient who was experiencing minor adverse events. It might, however, be possible to model the physician’s opinion by using either real or “paper” patients and providing the physician with a scenario (e.g., “if a new DMARD were available of equivalent efficacy and toxicity to MTX would you continue this patient’s current medication?”). This was the approach taken for developing the LDAS definition at OMERACT 7.

Alternatively, treatment response both in clinical practice and longitudinal observational studies might be defined in terms of longer term outcomes since the ultimate aim of DMARD therapy is to retard disease progression. It may be non-responders are those whose response is not adequate. Definitions of non-response and satisfactory response need to be developed separately, but ideally, using the same core outcome measures. The shortcomings of the currently used approaches have been highlighted. There are a number of options, however, that would justify further study.

One approach to define “satisfactory” response would be to model the physician’s decision to stop or continue the DMARD (in the same way as the DAS was modelled on the decision to start treatment). However, such decisions are context and patient specific. Thus, in an era when there was only one DMARD, patients would continue the drug for much longer than when there were another 6 to 8 DMARD to try. Similarly the individual patient may continue their 5th for longer than their 1st DMARD, even if the improvement observed is less, because there are fewer remaining options for treatment. In addition, drugs such as MTX are now being used at higher doses, so a patient considered to be a non-responder to MTX at 10 mg orally per week might respond to a higher (or parenteral) dose. The decision to continue a DMARD is also influenced by whether the patient is experiencing any side effects. Thus 2 patients may show a similar moderate response to a DMARD — and that DMARD would be continued in the patient who felt well, and discontinued in the patient who was experiencing minor adverse events. It might, however, be possible to model the physician’s opinion by using either real or “paper” patients and providing the physician with a scenario (e.g., “if a new DMARD were available of equivalent efficacy and toxicity to MTX would you continue this patient’s current medication?”). This was the approach taken for developing the LDAS definition at OMERACT 7.

Alternatively, treatment response both in clinical practice and longitudinal observational studies might be defined in terms of longer term outcomes since the ultimate aim of DMARD therapy is to retard disease progression. It may be that no DMARD (or combination of DMARD) can completely halt radiological progression especially in established RA. The important question then is whether, in terms of radiological outcome, drug response is truly a continuous spectrum: that is, that the lower the disease activity (measured using a continuous measure such as the DAS28), the less the radiological progression. If this is the case, then a definition of satisfactory response, beyond that of symptom control, is indeed arbitrary and depends on what proportion of the “tail” of Figure 1B the investigator is seeking to capture. Alternatively, there may be a threshold of disease activity below which there is no further influence on radiological (or disability) outcome. If this threshold could be identified, then the definition of “satisfactory response” could be based on a meaningful clinical construct.

All these response criteria apply to the assessment of primary response to the initiation of a new DMARD, seen in the first 6 to 12 months. They do not address the separate question of secondary non-response, seen in patients who initially respond to a DMARD that later becomes ineffective.

Finally it cannot be assumed that the DAS or DAS28 is the best composite measure of disease activity to develop definitions of drug response in patients with established RA. It may be that the same 4 components need to be combined using a different formula in this latter group of patients in whom the relationship between the number of inflamed joints and the ESR may not be the same as in early RA.

CONCLUSIONS
Treatment response to DMARD therapy is probably a continuous variable. It may or may not be normally distributed. Nevertheless clinicians categorize some patients as either non-responders or satisfactory responders to a particular drug. They are presumably recognizing the 2 tails of the distribution. However, many patients may be partial responders, although the decision about adding extra therapy depends both on patients’ and physicians’ expectations of treatment, together with the therapeutic options available. In the current era when there are several DMARD plus the biologics to choose from, it is important to be able to select a drug (or combination) to which the patient is likely to respond. Work on predicting response or non-response to a particular drug requires a robust and reproducible definition. However it may be appropriate to develop and test a number of definitions of response against the possible predictors.

Developing one or more definitions of primary non-response is likely to be the easiest task. They could be based on a minimum duration of exposure to the drug at an optimal dose plus the physician (and patient’s) decision to stop therapy (assuming that it was possible to disentangle whether the drug was stopped for inefficacy or side effects). Alternatively the ACR20 or DAS28 criteria could be used, but this would require measurement of the components of the criteria close to the time both of starting and stopping the
DMARD. Definitions of treatment response will depend on the perceived aim of DMARD therapy (symptom control vs slowing vs halting disease progression). The latter 2 definitions would require prospective followup of patients. In patients with long standing disease it may be difficult even to define adequate symptom control because existing joint damage may lead to pain, which cannot be modified by DMARD therapy.

Despite all the above reservations, the classifying of treatment response is a challenging area, but the problems are not insurmountable!

Samantha L. Hider, MSc, MRCP, Clinical Research Fellow; Ian N. Bruce, MD, Senior Lecturer and Honorary Consultant Rheumatologist; Alan J. Silman, MD, ARC Professor of Rheumatic Disease Epidemiology; Deborah P.M. Symmons, MD, Professor of Rheumatology and Musculoskeletal Epidemiology, ARC Epidemiology Unit, University of Manchester, Manchester UK.

Supported by the Arthritis Research Campaign UK. Address reprint requests to Dr. I.N. Bruce, ARC Epidemiology Unit, Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT UK. E-mail: ian@fs1.ser.man.ac.uk

REFERENCES