Remission of rheumatoid arthritis: should we care about definitions?

D. Aletaha¹, J.S. Smolen¹²

¹Department of Rheumatology, Medical University of Vienna, Austria; ²Second Department of Medicine, Hietzing Hospital, Vienna, Austria.

Daniel Aletaha, MD, MHS; Josef S. Smolen, MD.

Please address correspondence to:
Daniel Aletaha, MD, MHS, Department of Rheumatology, Internal Medicine III, Medical University of Vienna, Waehringer Guertel 18-20, A-1090 Vienna, Austria.
E-mail: daniel.aletaha@meduniwien.ac.at

Clin Exp Rheumatol 2006; 24 (Suppl. 43); S45-S51.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2006.

Key words: Rheumatoid arthritis, remission, disease activity.

ABSTRACT

A state of remission can be achieved in more and more rheumatoid arthritis (RA) patients. The combination of several RA disease activity measures seems to be important to provide an overall view of disease activity. Remission can be defined by two different approaches: one using a categorical model, requiring criteria for multiple variables to be fulfilled, each with its own threshold value (remission “criteria”); the other using a dimensional model, providing single measures of activity, which allow definition of remission by a single cut point (remission cut points for composite indices). The face validity of remission as defined by composite indices surpasses the one for the “criteria”. Likewise, the ones that are not weighted seem to surpass the weighted ones, as can be seen by the significant proportion of patients that continues to have considerable swollen joint counts despite being in Disease Activity Score (DAS)-28 remission. All composite indices seem to perform similarly well as tests for remission using expert judgments as the gold standard.

Quantifying disease activity in rheumatoid arthritis: How specific do we need to be?

With the advent of new therapies in the past decade, rheumatologists have become increasingly ambitious, and their new goal for treatment of rheumatoid arthritis (RA) is to completely avert the impact of this disease on their patients’ lives (1). As composite indices were developed to evaluate actual disease activity at every visit (2-5) rather than just response to therapy (6), these ambitions could be intensified to make even the smallest trace of rheumatoid arthritis (RA) disease activity measurable. The major advantages of these new measures are not only due to the continuous scale that they are based on. Their advantages are related especially to their ability to quantify the disease process even at the very low end of the activity scale, just as high-sensitivity assays for C-reactive protein (CRP) now allow quantifying the acute phase response at very low levels in RA patients with low disease activity and seemingly normal CRP or even in patients with osteoarthritis (7). Given these advances, outcomes assessment in RA is now heading towards a new problem, namely the recognition of our growing inability to clearly discern the presence of active disease from its absence. Clinically, these two states correspond to the terms “active disease” and “remission”. Therefore, if remission is the antonym of active disease, then the presence of one excludes the other, and in turn, the state of remission precludes any measurable disease activity.

One conclusion from the above paragraph is that the specificity of remission criteria is largely defined by the sensitivity of the activity measures at the bottom end of the spectrum. In addition, the classification of sustained remission heavily depends on how frequently disease activity is evaluated. However, the link between the specificity of remission criteria and the clinical need for being specific (as regards the potential sequelae of ongoing inflammation) still remains unclear. In this respect, two examples come to mind: some have argued that comprehensive joint counts should be used to be more specific for remission (8) suggesting that the residually involved joints not captured by reduced joint counts may have clinically meaningful long-term consequences. On the other hand, we have shown that exclusion of the acute phase response from the composite assessment of disease activity did not affect the performance...
compared to a score that includes such measure (5), which could be interpreted (although not primarily implied by us) in the sense that a residual systemic inflammatory response as displayed by an increase of CRP may not matter to long-term outcome of RA. In any case, the use of a composite index to assess RA disease activity seems to be inevitable from a clinical and methodologic perspective (9, 10), and none of these indices will capture all aspects of the disease should the index be practically useful and characterized by high sensitivity to change. Nevertheless, the issue with some composite indices is that it often becomes unclear what a given measured level on the respective scale means in terms of the underlying disease activity (11). Several questions are typically asked in this regard, which are fueled by the seemingly infinite heterogeneity of clinical presentations of patients with RA: Which elements of the disease are considered? Have those elements that need to be suppressed already become normal? Who are the stakeholders for identifying elements that need to be suppressed?

In the following section, we will briefly discuss some issues related to the perspectives of the different stakeholders and then look at the comparison of various remission criteria.

The patient’s perspective

In RA outcomes research, there have been ongoing efforts to include patients and their perspective into the assessments that are obtained in clinical practice and in trials. One argument to support such approach was that the perception of health status differs between patients and physicians (12). The respective activities at OMERACT revealed the importance of fatigue in the view of the patients, as well as the need for frequent disease activity assessment to accommodate temporal variations (13). In a study by Stamm et al., it was also seen that several aspects of the impact of the disease on patients are not necessarily well covered by current approaches (14). In another study, Heiberg and Kvien showed that patients gave the highest priority for improvement to their levels of pain (15).

In a similar way, we show here, on the basis of clinical observations, that patients consider themselves to experience higher levels of disease activity if tender joints are present as compared to swollen joints. We therefore used our observational cohort of patients with RA (16-18) at a cross-sectional point in time, excluding all patients for whom assessments were missing (n=767). We correlated the patient global assessment of disease activity (PGA) on a 0- to 10-cm visual analogue scale (VAS) with joint counts (on 28-joint scales). Correlation with tender joint counts was significantly better than correlation with swollen joint counts. In contrast, this was not the case when the evaluator global assessment of activity (EGA) was correlated with the 28-joint counts (Fig. 1A).

The physician’s perspective

The request that patients should decide which outcomes are relevant is important, because they are the ones affected by the disease. However, this can be challenged by several arguments. One is the fact that some aspects of the disease are harmful in the long term without causing pain or discomfort to the patient. It has, for example, been shown that swollen joint counts and the acute phase response are both associated with radiographic progression (5, 19-22). If no pain were present, patients would be willing to accept these symptoms and findings. In the same way as it would be unacceptable to guide treatment for hypertension purely based on patient reported outcomes (e.g., headache) as opposed to objective measures (i.e., the actual blood pressure) that are known to be closely related to preventable outcomes (i.e., cardiovascular events or renal failure), it is not acceptable for rheumatologists to base their treatment decisions on measures that are only patient derived. Nevertheless, some self-report measures, especially those assessing physical function, are still among the measures with the highest association with adverse long-term outcome. However, there are some issues related to guiding treatment decisions based on measures of physical function, which are discussed in a subsequent section (“remission and physical function”).

As mentioned above, in patients with RA, the self-assessment of disease activity is largely driven by pain (which by itself is a strong determinant of physical function), while other
symptoms that are also related to the inflammatory process, but are not painful, are prone to be neglected by the patient.

As an example, the data shown in Fig. 1A can be rearranged as presented in Fig. 1B, so that the effect of swollen joint counts and tender joint counts on the evaluation of disease activity by the different stakeholders can be explored: the number of swollen joints is closely correlated with the evaluator’s assessment of disease activity, and significantly less correlated with the patient’s assessment of disease activity. In contrast, tender joint counts have comparable correlations with both of the global scores (Fig. 2B). In other words: while physicians do not appear to neglect the feature deemed most important by the patients, patients do not appear to sufficiently appreciate the importance of a feature that, as physicians have learned, is more strongly associated with long-term outcome than the former.

Another piece of evidence stems from a recent survey, in which 44 international rheumatologists were asked to indicate the cut point for remission individually for the various core set variables (23). The physicians were first asked to consider a clinical situation where the variable of interest was the only one available to assess disease activity and then responded to the individual items on the survey (each corresponding to a single core set variable). The results showed that the median cut point for remission on a 28 joint scale was 1 for the swollen joint count and 2 for the tender joint count (Table I); this difference was based on a shift of the distribution curve, that is, it was seen at the 25th percentile in the same way as for the maximum response. In line with the findings mentioned above, this indicates that physicians consider swollen joints to be more reflective of an active disease process than tender joints.

Categorical and dimensional model of remission

It can be concluded from the above that it is one of the important responsibilities of the rheumatologist to guide treatment based on the interpretation of all available measures (subjective and objective) that are potential harbingers of a bad prognostic outcome. Using a set of different variables, the definition of a state like remission can principally be approached from two different perspectives, the categorical model and the dimensional model (24). In the categorical model, definition of the state requires fulfillment of multiple criteria (e.g., 5 of 6, such as in the Pinals criteria (25) for remission), where each has its own threshold value for being judged as positive or fulfilled. Categorical models are also typically used for classification criteria of various disorders. The dimensional model on the other hand provides an integrative view of the different elements that are taken into consideration. This is achieved by either creating scales that are based on a summation of categories.
of individual measures (e.g., the Glasgow Coma Scale) or by indices that summarize the actual values of different measures into a single number. Translated to rheumatology, the categorical model can be translated as an intersection of cut points for individual measures, whereas the dimensional model would best correspond to criteria for composite indices.

One important aspect in regard to the definition of clinical states, such as remission, is that the classification should ideally match the impression that rheumatologists would get in clinical practice. This is often called “face validity” of a definition. In this regard, the definition of remission based on composite indices is more likely to have overall face validity than definitions based on categorical criteria. In the latter, fulfilling a threshold value for one variable can compensate for fulfilling another criterion, thus potentially allowing some of the disease activity measures to be significantly elevated, as long as at least some of the others are normal. Composite indices overcome this problem, because they provide an integrative view of the included variables. Thus, each increment in one of the variables is associated with some increase in the overall value of the index. In the following discussion, we will focus on remission criteria for RA composite indices.

Remission criteria for RA composite indices
Currently, the most widely used composite indices are the Disease Activity Score (DAS) as originally described or as modified for 28 joint counts (DAS28) (2, 3), the Simplified Disease Activity Index (SDAI) (4), and the Clinical Disease Activity Index (CDAI) (5). These indices allow both evaluation of actual disease activity within or between individual patients or groups of patients and response to therapy (2-5, 11, 26). This applicability to reflect a disease activity state distinguishes these scores from the ACR response criteria. The latter need a given baseline state to which the current state can be compared to, and therefore, the actual disease activity upon fulfillment of these criteria can widely differ among individuals or patient samples (27, 28).

As a consequence of a reflection of the full range of disease activity in the composite scores, certain disease activity states can be categorized by derivation of cut points using “gold standards”. Obviously, such cut points and categories are at best approximations of the “true” state and may vary particularly due to differences in perception (i.e., changes within the gold standard) in the course of secular trends in time. The remission cut point for the original DAS is ≤1.6, while it had been defined for the DAS28 (based on a 28-joint count) as < 2.6. We have recently defined remission criteria based on a comprehensive survey among 35 international rheumatologists (29). These physicians were from 14 different countries, were experts in the field of RA clinical trials, and had a median time in rheumatology practice of 20 years. They were asked to indicate the presence of remission or active disease states on a set of 35 patient profiles that included all relevant clinical parameters. These responses served as the gold standard during the determination of the best cut points by means of various complex statistical procedures (29).

As one of the results, the cut point obtained for the DAS28 was similar, although somewhat lower (DAS28 ≤ 2.4) than the one originally defined. We also defined a remission cut point for the SDAI which was 11. In an additional analysis using the same methods, we derived the remission cut point for the CDAI as 10 (28).

Comparative validity of remission criteria for composite indices
We have discussed the potential advantages of indices over categorical criteria. However, even among the composite indices, there are differences in regard to face validity that stem from the different constructions of the scores. Fig. 2A-C illustrates this issue by comparing the cumulative distributions for seven disease activity measures in patients classified as being in clinical remission. One can see that, for example, about 85% of patients classified as in remission by the SDAI truly have no swollen joints, while this is the case only for 70% of patients classified as in DAS28 remission (Fig. 2A); also, residual joint counts in SDAI remission never exceeded 2, whereas they reached up to 10 and more joints in DAS28 remission (29-31). In addition, pain scores and patient and evaluator global scores showed cumulative remission distributions favoring the SDAI definition (Fig. 2B). Pain scores are of

---

**Table I. Indicated cut points for remission.**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Min.</th>
<th>10th pct.</th>
<th>25th pct.</th>
<th>50th pct.</th>
<th>75th pct.</th>
<th>90th pct.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swollen joint count (0-28 joints)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Tender joint count (0-28 joints)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Erythrocyte sedimentation rate (mm/h)*</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>14</td>
<td>19.5</td>
<td>25</td>
<td>30</td>
</tr>
<tr>
<td>C-reactive protein (mg/L)**</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>14.5</td>
<td>32</td>
</tr>
<tr>
<td>Patient assessment of activity (0-100 VAS)</td>
<td>0</td>
<td>4.5</td>
<td>7</td>
<td>12.5</td>
<td>19</td>
<td>24.5</td>
<td>39</td>
</tr>
<tr>
<td>MD assessment of activity (0-100 VAS)</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>8.5</td>
<td>12</td>
<td>20</td>
<td>37</td>
</tr>
<tr>
<td>Health assessment questionnaire disability index (0-3)</td>
<td>0</td>
<td>0.1</td>
<td>0.3</td>
<td>0.5</td>
<td>0.5</td>
<td>1.1</td>
<td>1.5</td>
</tr>
</tbody>
</table>

*All cut points can be understood as “≤” (for details, see text)

† If the indicated ESR was < 2 mm, then ESR was set to 2 mm (suggested lowest possible ESR value)

‡ If the indicated CRP was < 5 mg/L then CRP was set to 5 mg/L (lower limit of detection for most routine laboratories) for all responses <5 mm on the VAS.

Adapted from Ref. (23).
particular interest in this comparison, since neither of these measures is included in the SDAI or DAS28 formulae. Erythrocyte sedimentation rate (ESR) is the variable that is most stringently represented in the DAS28 (Fig. 2C). Another aspect is construct validity, and this relates to important outcomes that make sense knowing the nature of the disease. In the case of RA, functional deterioration and radiographic progression are two measures of construct validity. After derivation of the remission criteria for the SDAI and the DAS28 (29), we analyzed how patients classified into this state would progress radiographically over time. In Fig. 3, it can be seen that patients who spend the majority of a given period of time (in this case, 3 years) in remission, as defined by the mentioned criteria for the DAS28 (Fig. 3, left panel) or the SDAI (Fig. 3 right panel), do not progress radiographically on average, and significantly less than those who spent 50% or less time in remission (29).

Finally, the overall value of criteria in classifying a subject can be assessed using diagnostic testing procedures. In receiver operating characteristics (ROC) curve analysis, a series of increasing cut points on a continuous scale (in our case, SDAI or DAS28 scale) are used to classify a set of patients, and sensitivity and specificity of this classification is compared to a gold standard classification. We used the mentioned physician’s classification as the gold standard and performed a ROC curve analysis on the SDAI, CDAI, and DAS28. The area under the ROC curve is a measure of the overall usefulness of a scale as a test for a state and should be significantly different from 0.5, which is the result for a useless test. From Table II, it can be seen that the present remission criteria conform very well with the view of international experts in the field of RA, with area under the ROC curves of >0.94. Although in some circumstances the SDAI conforms better with physicians’ decision making than other scores (32), one can conclude from these results that there is overall no significant difference between these scores in regards to their accuracy of classification, as had also been shown for other disease activity states based on these indices (33). Nevertheless, the frequencies of classifying patients as in remission are higher for the DAS28 than for the SDAI and CDAI (34), which occurs because patients can be classified as in DAS28, but not SDAI or CDAI remission even if they have a considerable residual swollen joint count and pain (see above and Fig. 2). In view of the association of swollen joint counts with progression of joint destruction (35), it will have to be seen if these patients, whose other components of the DAS28 have reached a level low enough to fulfill remission criteria, experience higher degrees of joint damage and disability in the long term.

Remission and functional status
Disability is the major sequela of RA, and under ideal circumstances remission will be associated with normalization of functional impairment as measured by the Health assessment questionnaire (HAQ). However, in contrast to most other variables used for disease activity assessment, such as joint swelling or acute phase responses, disability is not monomorphic. Rather, it is dependent on various factors. In RA, disability has two major components, one being a consequence of active disease, and the other related to joint damage (36); these two components create a duality of RA disability with reversible elements on the one hand, and irreversible elements on the other. The irreversible component of disability increases with increasing joint damage and increasing disease duration (37). Moreover, even in the general healthy population, disability increases with increasing age (38). Therefore, when using functional instruments to assess remission, potential cut points would not be generalizable to the whole population of RA patients; rather several cut points, each for a particular subgroup

Table II. Usefulness of composite indices as tests for remission.

<table>
<thead>
<tr>
<th>Test result variables</th>
<th>Area under the ROC curve(^{(95% \text{ confidence interval})})</th>
<th>(p)-value(^{3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplified Disease Activity Index (SDAI)</td>
<td>0.952 (0.871 to 1.033)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Clinical Disease Activity Index (CDAI)</td>
<td>0.946 (0.854 to 1.037)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Disease Activity Score 28 (DAS28)</td>
<td>0.942 (0.863 to 1.021)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

\(^{3}\) Receiver operating characteristics (ROC) curve analysis using the classification of remission by the majority of 35 rheumatologists as the gold standard for remission.

\(^{3}\) \(p\)-value for test using the null hypothesis of an area under the curve (AUC) = 0.5.
of patients, would be needed. Even then a given HAQ score would not inform if a patient has no evidence of active disease but had gone through significant joint destruction or has no damage but residual disease activity. In fact, improvement of HAQ scores from baseline to clinical (SDAI) remission differed significantly depending on the degree of joint damage: while about 25% of patients with moderate to severe joint damage had less than 50% improvement, more than 75% of patients with little joint damage had more than 80% improvement of the HAQ (37).

In this context the data shown in Fig. 4 are of interest, since they relate HAQ scores to disease activity states obtained by two composite activity indices, the SDAI and the DAS28. Patients classified as being in low disease activity by both SDAI and DAS28 had virtually identical median HAQ values, as well as HAQ distributions. In contrast, in remission by the SDAI, the 75th percentile of the HAQ values did not exceed 0.5 whereas it extended to 0.75 among patients in DAS28 remission; likewise, the median HAQ tended to be higher in DAS28 than SDAI remission (Fig. 4). This finding is in line with the higher joint counts observed in DAS28 remission as discussed above.

Thus, while functional status is an excellent predictor of mortality in RA (39), function is not an ideal attribute to guide treatment decisions, since, in contrast to pure measures of disease activity, two patients with similar HAQ scores might be orthogonally different in regards to the required therapeutic consequence. Therefore, we feel that measures of function should be used separately from other instruments of disease activity assessment in the evaluation of remission, at least until ways are found to interpret the reversible and irreversible components of disability separately; such investigations are on the way. In any case, since measures of function and measures of pure disease activity convey different, though partly overlapping information, functional assessment should definitely be part and parcel of RA evaluation, including the assessment of remission.

Conclusion

Remission is a state that can be achieved in an increasing proportion of patients. The combination of several RA disease activity measures seems to be important for an overall view of disease activity. Remission can be defined by two different approaches: one using a categorical model, requiring criteria for multiple variables to be fulfilled, each with its own threshold value (remission "criteria"); the other using a dimensional model, providing single measures of activity, which allow definition of remission by a single cut point (remission cut points for composite indices). The face validity of remission as defined by composite indices surpasses the one for the "criterion", and among the composite indices, the ones that are not weighted seem to surpass the weighted ones. All composite indices seem to perform similarly well as tests for remission using expert judgments as the gold standard, although in DAS28 remission a significant proportion of patients continues to have considerable swollen joint counts and HAQ scores.

References

9. Tugwell P, Bombardier C: A methodologic framework for developing and selecting...
18. KRAPL T, STAMM T, MACHOLD KP, MONTAG K, SMOLEN A, ALETAHA D: Methotrexate in rheumatoid arthritis is frequently effective, even if re-employed after a previous failure. *Arthritis Res Ther* 2006; 8: R46.

---

**Should we care about RA remission criteria? / D. Aletaha & J.S. Smolen**

---