The Problem of Rheumatoid Arthritis Disease Activity and Remission in Clinical Practice

TIMOTHY S. SHAVER, JAMES D. ANDERSON, DAVID N. WEIDENSAUL, SHADI S. SHAHOURI, RUTH E. BUSCH, TED R. MIKULS, KALEB MICHAUD, and FREDERICK WOLFE

ABSTRACT. Objective. To investigate the results and feasibility of available scales to measure minimal disease activity (MDA) and remission in rheumatoid arthritis (RA) in the clinic.

Methods. We studied 849 consecutive patients with RA seen in a community rheumatology practice for routine RA care by 4 rheumatologists, beginning in March 2007 and ending in August 2007. Patients and physicians completed a simple form at each visit. We calculated the Disease Activity Score 28 (DAS28), Clinical Disease Activity Index (CDAI), physician assessment of global activity, and the Patient Activity Scale (PAS-II). From these we calculated remission and MDA prevalence in this community practice.

Results. The DAS28 could not be determined in more than 50% of patients because of referring physician and insurance company restrictions. Remission prevalences differed by assessment method: DAS28 28.5%, CDAI 6.5%–8.1%, physician global 12.5%, PAS 13.7%. MDA was 26.9% using the American College of Rheumatology core set variables, 34.7% using the DAS28, and 26.8% using the PAS-II. The kappa statistic was only fair (0.2 to 0.4) for most comparisons between assessment methods. No significant differences were noted for remission and MDA according to biologic therapy.

Conclusion. The CDAI and/or physician global and PAS-II are simple acceptable ways to measure RA activity in the clinic, but results differ strikingly according to method. Further standardization appears to be required for full implementation of the CDAI. Caution is urged before using these methods for regulatory purposes. (First Release April 15 2008; J Rheumatol 2008;35:1015–22)

Key Indexing Terms:
RHEUMATOID ARTHRITIS ASSESSMENT CLINICAL DISEASE ACTIVITY INDEX REMISSION MINIMAL DISEASE ACTIVITY DISEASE ACTIVITY SCORE
series of measures “physician measures” and the second series “patient measures.”

Measurement has taken on a new imperative as regulators and insurance companies impose requirements on use of expensive therapies. Yet from the clinicians’ perspective, physician and patient measures may be overkill, consume valuable time, and be expensive, and may not always be available. For example, if a patient’s activity and symptoms are stable, what is to be gained by the clinician by detailed joint and laboratory examinations?

We investigated several measures of remission and disease activity that address either physician or patient measures or both. We investigated the prevalence of remission and minimal disease activity (MDA) in all patients and those receiving and not receiving biologics, the agreement between measures, the ability to obtain complete scales, and factors associated with scale differences.

MATERIALS AND METHODS

We studied 849 consecutive patients with RA seen in a community rheumatology practice [Arthritis and Rheumatology Clinics of Kansas (ARCK)] for routine RA care by one of 4 rheumatologists, beginning in March 2007 and ending in August 2007 (5 months). If a patient was seen more than once during the study period the most recent encounter was analyzed. Each patient was asked to complete a self-report questionnaire that consisted of the 10-item HAQ-II functional scale, a 5-category health satisfaction question, and visual analog scales (VAS) for pain, fatigue, and global severity. The reverse side of the questionnaire was completed by the rheumatologist and consisted of tender and swollen joint manikins, a medication list, a section for reporting ESR or CRP, and physician’s assessment of global disease activity. From the joint manikins we determined the 28 swollen and tender joint counts. The physician’s global question was an 11-item rating index of the rheumatologist, including vasculitis, scleritis, etc. Clinic staff were instructed on how to keep records and how to ensure that patients and physicians completed all questions. A followup meeting 1 week later identified problem areas. This led to modifications of the study routine and questionnaire. Following completion of the questionnaires, the questionnaires were sent to the National Data Bank for Rheumatic Diseases (NDB) database, Wichita, KS, where they were scanned into the NDB data base. Omissions and errors were identified immediately and immediate feedback was given to ARCK. After a few days of interaction, omissions and errors became minimal.

Remission indexes. In our study, we use the word “remission” to mean satisfying remission criteria without consideration of the duration of remission, as our purpose was to examine the level of disease activity necessary to define remission.

We calculated the Disease Activity Index 28 (DAS28) from the swollen and tender joint count, the ESR, and patient’s VAS global pain and fatigue in all patients and those receiving and not receiving biologics, the agreement between measures, the ability to obtain complete scales, and factors associated with scale differences.

Minimal disease activity (MDA). We calculated 3 indexes: the MDA Core Set Index was positive if the patient satisfied at least 5 of 7 of the following conditions: VAS pain ≤ 2, swollen joint count ≤ 1, tender joint count ≤ 1, HAQ ≤ 0.5, patient global severity ≤ 2, physician global disease activity ≤ 1.5, ESR ≤ 20 mm/h; OR patient satisfied the following conditions: the patient had no swollen joints, no tender joints, and an ESR ≤ 10 mm/h. The definition was called the “Core Set” definition because it made use of the published core set variables for the assessment of response in RA. In the second method, “DAS28 method,” patients achieved MDA status if the patient had no swollen joints, no tender joints, and an ESR ≤ 10 mm/h OR the DAS28 score was ≤ 2.85. In the third method, PAS MDA, a patient was MDA-positive if PAS-II score was ≤ 2.218. This value was modified from an original PAS value of 1.75. The PAS-II is a summary measure of RA activity made by combining HAQ-II2 and VAS pain and patient global scales. The range is 0–10, with higher values indicating more abnormality. The MDA Core Set and DAS28 methods were developed as part of the Outcome Measures in Rheumatology [Clinical Trials] 7 process.

To assess the effect of methodology on remission and MDA as a function of biologic therapy, we characterized a patient as being on a biologic if he was receiving infliximab, etanercept, adalimumab, anakinra, abatacept, or rituximab.

Statistics. We separately assessed agreement between both remission and MDA indices by the kappa statistic. Fleiss indicates the interpretation of the kappa statistic in terms of strength of agreement as < 0.00 poor, 0.00–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect. Differences between physician examinations and ratings were analyzed by linear and logistic regression. Figure 1 was created using Lowess regression. Data were analyzed using Stata (Stata, College Station, TX, USA) version 10.0. Statistical significance was set at the 0.05 level, confidence intervals were established at 95%, and all tests were 2-tailed.

RESULTS

The mean age of patients with RA was 59.1 (SD 13.6) years and 24.8% were men. The median duration of RA was 7.2 years, and 16.6% had RA for 2 years or less.

Table 1 shows percentages, means, and percentage missing data for a series of remission, MDA, and disease activity measures. Missing data were primarily determined by missing ESR values (46.8%), and this in turn was responsible for the inability to collect complete data for the DAS28, where the missing data level was 48.8%. According to ARCK clinical physicians, missing ESR data was caused by failure of primary care physicians, insurers, and health maintenance organizations to authorize or provide the ESR. The CDAI, which primarily requires physician data (3 items) and patient data (1 item), was missing in only 6.6%, the MD global remission was missing in 3.2%, and the PAS (all patient data) was missing in 5.1%.
Figure 1. Plot of DAS28 values against CDAI values (n = 435). The line is determined by Lowess regression. Circles below the horizontal line at 2.6 represent patients in DAS28 remission. Circles to the left of the vertical line at 2.8 represent patients in CDAI remission. Kappa for the DAS28 and CDAI scales was fair, 0.2536.

Table 1. Remission, minimal disease activity, and clinical variables in 849 patients with RA.

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Data (n = 435)</th>
<th>Complete Data (n = 408)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>% Missing</td>
</tr>
<tr>
<td>Remission indices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS28 ≤ 2.6</td>
<td>435</td>
<td>48.8</td>
</tr>
<tr>
<td>CDAI ≤ 2.8</td>
<td>795</td>
<td>6.6</td>
</tr>
<tr>
<td>CDAI ≤ 3</td>
<td>795</td>
<td>6.6</td>
</tr>
<tr>
<td>MD global ≤ 1</td>
<td>816</td>
<td>4.9</td>
</tr>
<tr>
<td>PAS-II ≤ 1.25</td>
<td>822</td>
<td>3.2</td>
</tr>
<tr>
<td>Swollen, tender joints = 0</td>
<td>847</td>
<td>1.4</td>
</tr>
<tr>
<td>Swollen, tender joints = 0, ESR &lt; 10</td>
<td>453</td>
<td>46.8</td>
</tr>
<tr>
<td>Minimal disease activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS28 criteria (DAS28 ≤ 2.85)</td>
<td>435</td>
<td>48.8</td>
</tr>
<tr>
<td>Core criteria</td>
<td>424</td>
<td>50.1</td>
</tr>
<tr>
<td>PAS-II ≤ 2.2</td>
<td>822</td>
<td>3.2</td>
</tr>
<tr>
<td>Clinical variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS28 score (0–8.1)</td>
<td>435</td>
<td>48.8</td>
</tr>
<tr>
<td>CDAI (0–76)</td>
<td>795</td>
<td>6.6</td>
</tr>
<tr>
<td>PAS-II (0–10)</td>
<td>822</td>
<td>3.2</td>
</tr>
<tr>
<td>ESR, mm/h</td>
<td>453</td>
<td>46.8</td>
</tr>
<tr>
<td>MD 28 swollen joint count (0–28)</td>
<td>847</td>
<td>1.4</td>
</tr>
<tr>
<td>MD 28 tender joint count (0–28)</td>
<td>847</td>
<td>1.4</td>
</tr>
<tr>
<td>MD global (0–10)</td>
<td>816</td>
<td>4.9</td>
</tr>
<tr>
<td>Pain (0–10)</td>
<td>849</td>
<td>0.0</td>
</tr>
<tr>
<td>Global severity (0–10)</td>
<td>847</td>
<td>0.5</td>
</tr>
<tr>
<td>HAQ-II (0–3)</td>
<td>829</td>
<td>2.7</td>
</tr>
</tbody>
</table>

* Percentage positive excluding missing observations. DAS: Disease Activity Score; CDAI: Clinical Disease Activity Index; PAS: Patient Activity Scale; ESR: erythrocyte sedimentation rate; HAQ: Health Assessment Questionnaire.
With respect to remission, prevalence was quite different depending on the index examined (Table 1). DAS28 remission occurred in 28.5%, PAS-II remission in 13.7%, MD remission in 12.5%, and CDAI remission in 6.5%. When a cutoff of 3.0 was used (to adjust for slight scale differences) for CDAI, the remission proportion was 8.1%. MDA was 26.9% with Core criteria, 34.7% with DAS-28 criteria, and 26.8% with PAS-II criteria. There was little difference in prevalences for remission and MDA for patients with complete or incomplete data (Table 1).

We next investigated agreement among remission scales (Table 2, Figure 1). Agreement between most established scales was fair (kappa 0.2–0.4), but was slightly better (moderate) for CDAI and MD global remission. Similar prevalences, therefore, do not mean the same patients are in remission or MDA. Although the kappa for agreement between DAS28 remission and CDAI remission was only 0.253, the Pearson correlation between DAS28 and CDAI as full scales was 0.826.

We further explored prevalence differences for the DAS28 (28.5%) and CDAI (6.5%) by examining the characteristics of patients in DAS28 remission but not in CDAI remission (lower right quadrant of Figure 1) in Figure 2 and by examining the characteristics of patients in CDAI remis-

Table 2. Kappa scores for paired remission categories

<table>
<thead>
<tr>
<th></th>
<th>DAS28 Remission (DAS28 ≤ 2.6)</th>
<th>Swollen, tender joints = 0</th>
<th>PAS-II Remission (PAS ≤ 1.25)</th>
<th>MD Remission (MD global ≤ 1)</th>
<th>CDAI Remission</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swollen, tender joints = 0</td>
<td>0.330</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAS-II remission</td>
<td>0.252</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAS ≤ 1.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD remission</td>
<td>0.318</td>
<td>0.428</td>
<td></td>
<td></td>
<td>0.286</td>
</tr>
<tr>
<td>(MD global ≤ 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDAI remission</td>
<td>0.253</td>
<td>0.321</td>
<td>0.404</td>
<td>0.506</td>
<td></td>
</tr>
<tr>
<td>Swollen tender joints = 0, ESR &lt; 10</td>
<td>0.366</td>
<td>0.431</td>
<td>0.115</td>
<td>0.297</td>
<td>0.345</td>
</tr>
</tbody>
</table>

DAS28: Disease Activity Score 28 joint count; PAS: Patient Activity Scale; CDAI: Clinical Disease Activity Index; ESR: erythrocyte sedimentation rate.

Figure 2. Histograms of CDAI remission variables in patients who are DAS28 remission-positive and CDAI remission-negative (n = 100).
sion (Figure 3) regardless of DAS28 score. Many patients in DAS28 but not CDAI remission have high patient global scores and more than minimal MD global scores, and some patients have high joint count scores (Figure 2). By contrast, CDAI remission patients have generally low scores for all measures (Figure 3).

To obtain a clearer sense of the relation between physician rating of disease activity and the DAS28 and CDAI scale, we performed logistic regression in 2 steps. First, we regressed MD remission (MD global ≤ 1) on DAS28 remission and CDAI remission. The odds ratio (OR) for DAS28 remission was 4.1 (95% CI 2.1–8.0) and the OR for CDAI remission was 31.6 (95% CI 10.0–100.2); the OR for the difference between these remission scores was 7.7 (95% CI 1.9–32.1). We also divided patients into 4 groups and determined the OR for the association of MD remission: no remission by any criteria (reference group); DAS28 remission but not CDAI remission, OR 4.2 (95% CI 2.1–8.6); CDAI remission but not DAS28 remission, OR 49.1 (95% CI 4.8–497.0); and remission by both criteria, OR 114.5 (95% CI 31.0–422.1). Although stronger associations with CDAI were to be expected because CDAI contains physician global, these results have important implications, given that the physician is the ultimate arbiter of disease activity in the setting of patient care.

We also studied whether differences between examining physician’s evaluations might alter CDAI and DAS28 results. In this model we controlled for severity by including age, sex, patient global, patient pain, ESR, and RA duration as covariates. We then regressed separately swollen joint counts, tender joint counts, and physician global on individual examining physicians and the covariates noted above. Individual physicians differed in their assessments of global severity (p < 0.001), tender joint counts (p = 0.023), and swollen joint counts (p < 0.001). We also tested physician differences by additionally adding as covariates the 2 omitted variables to the model from the 3 physician variables (swollen joint counts, tender joint counts, and physician global). Statistical significance was maintained in all analyses. These data show that physicians in our study systematically differed from each other with respect to their results for swollen joint counts, tender joint counts, and physician global.

Finally, we noted that use of biologics was not associated with more than minor differences in remission and MDA percentages for the key variables, DAS28, CDAI, physician’s global, and PAS-II measures (Table 3).

DISCUSSION
One of the important objectives of our study was to examine the utility of the study measures in clinical practice. The DAS28 is a widely used and well documented scale. However, our data show that it was possible to obtain DAS28 scores in fewer than 50% of patients because of issues involving non-allowance of the test by referring primary care physicians and insurers, as reported by ARCK clinic physicians. Although rheumatology experts might decry this result, it can be observed that a person with complete data (ESR, in particular) did not seem to have other scores that were different from those with incomplete data.

Figure 3. CDAI remission variables in patients who are CDAI remission-positive (n = 28).
It might be argued, further, that ESR testing is not needed or not needed as frequently. Our data cannot answer such a question, but they do raise the issue.

The inability to obtain ESR tests or to obtain them in a timely fashion so they could be used for clinical decision-making was one of the reasons that the CDAI was developed. The CDAI, in our hands, was simple to calculate and easy to use. However, we found that the remission prevalence using the DAS28 was 3 to 4 times greater than with the CDAI (Table 1, Figure 1). These results, CDAI remission prevalence of 6.5% to 8.1%, are different from those published by the CDAI developers, who found remission in 33.5% of patients with CDAI and 42.7% by DAS28, and led us to examine the issue further. To do this we examined 2 external sources. First, we used the RAES database. RAES was an observational study in which 644 patients were evaluated by DAS28 and CDAI by 61 physicians. As shown in Figure 4, the same pattern of reduced CDAI remission compared with DAS28 remission that we found in our study (Figure 1) was noted in the RAES database, and the remission prevalence was 7.8%. We also examined the results of the CDAI and DAS28 tests in more than 5,000 patients in Europe who participated in the QuestRA project and a separate cohort of over 800 US veterans with RA participating in the V ARA Registry, and found the same pattern (personal communication from Tuulikki Sokka, MD, and Ted R. Mikuls, MD, MSPH, respectively; manuscripts in preparation).

Figures 2 and 3 provide insight into the remission differences of the CDAI and DAS28 scales. Many patients satisfying DAS28 remission criteria but not CDAI remission had high patient global scores and more than minimal MD global scores, and some patients have high joint count scores (Figure 2). By the nature of the CDAI definition, CDAI remission patients generally had low scores for all of the incorporated measures (Figure 3). In addition, there was greater agreement between physician global (and physician remission) with the CDAI compared with the DAS28.
Although that was to be expected given that the physician global is part of the CDAI scale. The more stringent CDAI criteria had been noted previously by the CDAI authors\(^3,6\), who found 33.5% of patients in remission by CDAI and 42.7% by DAS28 in a clinical practice setting\(^3\). Their prevalence for CDAI remission was 3 to 4 times greater than in our study and the RAES database.

The remission percentages shown in Table 1 for a series of different possible scales illustrate the problem of remission ascertainment when remission depends strongly on the remission definition. We note with interest that the remission percentage based on the physician’s global (12.5%) is closer to the CDAI (6.5%) or the CDAI at a cutpoint of 3.0 (8.1%) than with the DAS28 (28.5%). Finnish investigators have also reported that the DAS28 remission cutoff may be too high\(^8\). Using the American College of Rheumatology preliminary criteria for remission\(^15\), they indicated that the best cutpoint for the DAS28 was 2.32. However, they noted that even with that cutpoint, a “substantial proportion of patients below the DAS28 cut off point for remission have tender or swollen joints, or both. DAS28 may not be an appropriate tool for assessment of remission in RA.” A similar residua of active joints for patients in DAS28 remission was noted by the CDAI developers\(^3\). In our study, using a DAS28 cutpoint of 2.32, 20.3% of patients with RA met that criterion. Both the DAS28 and CDAI criteria also run up against the problem that they do not include the hips and feet, thereby potentially missing evidence of RA activity and misclassifying remission.

Therefore, there are problems with any set of current remission criteria, and this can be further illustrated with clinic data. If a physician determines that a patient is in remission, but that patient has a high tender joint count (for example, 3 tender joints for the CDAI) or a high ESR for reasons that are not related to RA activity, that patient cannot be designated as being in remission when criteria, but not clinical judgment, are used. This discrepancy, which also can go in the other direction, illustrates the differences between criteria that can be used in studies of groups of patients and criteria that are appropriate and useful for the single patient in clinical practice.

In addition to the group versus individual patient difficulty, Altman and Royston indicate that in dichotomizing continuous states “...one may seriously underestimate the extent of variation in outcome between groups, such as the risk of some event, and considerable variability may be subsumed within each group. Individuals close to but on opposite sides of the cutpoint are characterised as being very different rather than very similar”\(^29\).

Additional problems with using measures such as the DAS28 and CDAI in the clinic relate to poor reliability. In our study, the 4 physicians differed in their assessments of joint swelling and tenderness and in physician global. It is known that these measures have poor reliabilities (< 0.80), and the DAS has a reliability around 0.80\(^22,30,31\). From reliability we can estimate the minimal detectable change (MDC; also called the reliable change or the smallest real difference)\(^32\)

\[
MDC = SEM \times 1.96 \times \sqrt{2}
\]

where \(SEM = SD \times \sqrt{1 - \text{reliability}}\)

The consequence of this is that, given 2 measurements in an individual patient, we can only say with confidence that differences between 2 assessments that are equal to or exceed ~2.0 for DAS28 are significant. The reason that the necessary differences are so great is that the uncertainty (reliability estimate) is applied not to the change score but to each of the 2 test measurements\(^33\).

Another approach to RA activity is found in the MDA method. This method avoids the large differences between measures that might be seen because of examination differences at the margin of remission. The MDA for DAS was 34.7%, for the Core measures was 26.9%, and for the PAS-II was 26.8%, in general agreement with each other. We have not attempted to define MDA for CDAI.

While we have indicated that the CDAI is more likely to be collected than the DAS28 because of problems obtaining the ESR, it should not be concluded that the CDAI is optimal for clinical practice. The physicians in our study were highly motivated with respect to joint counts, and were provided with ongoing feedback by the NDB quality control staff. It is likely that the percentage of missing joint counts or other physician data will be much greater in actual practice. One potential solution to this problem is to use patient data, since their missing proportion is likely to be consistently low. However, the correlation between patient and physician data is only fair, with kappas between 0.252 and 0.402. It should also be noted that the difficulty obtaining ESR noted in our study might not be the case in other settings.

If one is going to collect physician data, however, the CDAI is one of the simplest and most intuitive measures available, and it would allow complete data collection among the patients in our study. We would also observe that the addition of the HAQ-II and a pain scale would allow complete assessment of both patient and physician data. Clinicians using such data would have full documentation of their patients’ courses. For physicians who do not want to do formal joint examination, the physician global and the components of PAS-II allow rapid assessment. The physician global is performed by every physician, although not necessarily written down, but it is easy enough to do so.

The results of our study, showing wide differences in remission percentages, indicate a need for a general consensus on remission criteria definition. They also indicate the hazard of blindly using remission criteria in the clinic (Figures 1, 4). Use of CDAI scores on a continuous scale as
a guide, combined with the experienced clinician’s judgment, may be an appropriate path to assessing remission in the clinic.\(^4\)

The prevalence of remission differs strikingly according to definition and scales. The DAS28 was available in < 50% because of incompletely available ESR values, suggesting the CDAI might be more clinically useful. However, additional studies are required to determine appropriate levels for disease activity and remission. The physician’s determination of remission, considering all available evidence, may be the best approach to remission in the clinic. However, even when patients are similar, examination differences may lead to variable results.

ACKNOWLEDGMENT

We thank Dr. Tuulikki Sokka for making her QuestRA data available for our review, and for her thoughtful and very helpful comments on the manuscript.

REFERENCES


5. Landewe R, van der Heijde D, van der Linden S, Boers M. Twenty-eight-joint counts invalidate the DAS28 remission definition owing to the omission of the lower extremity joints: a comparison with the original DAS remission. Ann Rheum Dis 2006;65:637-41.


34. van Riel PL, Fransen J. To be in remission or not: is that the question? Ann Rheum Dis 2005;64:1389-90.