Validity and Reliability Problems with Patient Global as a Component of the ACR/EULAR Remission Criteria as Used in Clinical Practice

KARIM R. MASRI, TIMOTHY S. SHAVER, SHADI H. SHAHOURI, SHIRLEY WANG, JAMES D. ANDERSON, RUTH E. BUSCH, KALEB MICHAUD, TED R. MIKULS, LIRON CAPLAN, and FREDERICK WOLFE

ABSTRACT. ***Objective.*** To investigate what factors influence patient global health assessment (PtGlobal), and how those factors and the reliability of PtGlobal affect the rate, reliability, and validity of recently published American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) rheumatoid arthritis (RA) remission criteria when used in clinical practice.

Methods. We examined consecutive patients with RA in clinical practice and identified 77 who met ACR/EULAR joint criteria for remission (≤ 1 swollen joint and ≤ 1 tender joint). We evaluated factors associated with a PtGlobal > 1, because a PtGlobal ≤ 1 defined ACR/EULAR remission in this group of patients who had already met ACR/EULAR joint criteria.

Results. Of the 77 patients examined, only 17 (22.1%) had PtGlobal ≤ 1 and thus fully satisfied ACR/EULAR criteria. A large proportion of patients not in remission by ACR/EULAR criteria had high PtGlobal related to noninflammatory issues, including low back pain, fatigue, and functional limitations, and a number of patients clustered in the range of PtGlobal > 1 and ≤ 2. However, the minimal detectable difference for PtGlobal was 2.3. In addition, compared with a PtGlobal severity score, a PtGlobal activity score was 3.3% less likely to be abnormal (> 1).

Conclusion. Noninflammatory factors contribute to the level of PtGlobal and result in the exclusion of many patients who would otherwise be in "true" remission according to the ACR/EULAR definition. Reliability problems associated with PtGlobal can also result in misclassification, and may explain the observation of low longterm remission rates in RA. As currently constituted, the use of the ACR/EULAR remission criteria in clinical practice appears to be problematic. (First Release May 15 2012; J Rheumatol 2012;39:1139–45; doi:10.3899/jrheum.111543)

Key Indexing Terms:
RHEUMATOID ARTHRITIS REMISSION RELIABILITY

From the University of Kansas School of Medicine, the Arthritis and Rheumatology Clinics of Kansas, Wichita State University, and the US National Data Bank for Rheumatic Diseases, Wichita, Kansas; University of Nebraska Medical Center, Omaha VA Medical Center and University of Nebraska, Omaha, Nebraska; and Denver VA Medical Center and University of Colorado, Denver, Colorado, USA.

Dr. Caplan is supported by VA HSR&D CDA 07-221. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the US Department of Veterans Affairs or the US government.

K.R. Masri, MD, University of Kansas School of Medicine; T.S. Shaver, MD; S.H. Shahouri, MD; S. Wang, MD; J.D. Anderson, MD, Arthritis and Rheumatology Clinics of Kansas, University of Kansas School of Medicine; R.E. Busch, ARNP, Arthritis and Rheumatology Clinics of Kansas, Wichita State University; K. Michaud, PhD, University of Nebraska Medical Center, National Data Bank for Rheumatic Diseases; T.R. Mikuls, MD, MSPH, Omaha VA Medical Center and University of Nebraska; L. Caplan, MD, PhD, Denver VA Medical Center and University of Colorado; F. Wolfe, MD, National Data Bank for Rheumatic Diseases, University of Kansas School of Medicine.

Address correspondence to Dr. F. Wolfe, National Data Bank for Rheumatic Diseases, 1035 N. Emporia, Suite 288, Wichita, KS 67214, USA. E-mail: fwolfe@arthritis-research.org

Accepted for publication February 17, 2012.

In 2011, the American College of Rheumatology and the European League Against Rheumatism (ACR/EULAR) established a "Preliminary Definition of Remission in Rheumatoid Arthritis for Clinical Trials"[1], in which the authors also suggested "that a definition of remission be developed for clinic based practice that would not require an acute phase reactant, as long as it would capture remission as stringently as the measure employed for clinical trials," and further that "...core set measures should be used to define remission and that any definition of remission in clinical trials should look toward and make possible a similar definition in clinical practice." The ACR/EULAR authors then cautiously proposed clinical practice criteria for clinical settings where acute-phase reactants might not be available, but noted that the clinical practice criteria had not yet been validated. These 3 clinical criteria definitions included (1) ≤ 1 swollen joint, ≤ 1 tender joint, patient global (PtGlobal) ≤ 1 (ACR/EULAR criteria); (2) ≤ 1 swollen joint, ≤ 1 tender joint, PtGlobal ≤ 1, physician global score < 1 [ACR/

EULAR-4 variable criteria (AE-4)]; and (3) Clinical Disease Activity Index (CDAI) ≤ 2.8[2].

We recently evaluated all ACR/EULAR criteria in a large longitudinal study in clinical practice[3], and noted that the 3 definitions yielded cross-sectional remission rates between 5% and 10%, but that a remission lasting 2 years occurred in < 3% of patients. In trying to understand the lack of remission durability, we observed that at the initial time when remission would otherwise be achieved, many patients demonstrated a PtGlobal close to the remission level, but failed to satisfy the formal remission definition. Further, at subsequent clinic visits many patients previously in remission no longer met the definition of remission, primarily because of changes in PtGlobal. We also noted that PtGlobal was more strongly associated with remission than any of the other component remission variables (Appendix 1[3]). We concluded that "...remission in this setting [the clinic] depends on PtGlobal, which may reflect true sensitivity to changes in RA activity or represent reliability issues where remission status changes while RA activity actually remains the same."

To our knowledge, there are few studies that identify the factors that contribute to the level of PtGlobal in RA[4], and none in the context of remission. In our current study, we investigated what factors influence PtGlobal, and how these factors and the reliability of the PtGlobal might influence the rate of ACR/EULAR remission. To perform this evaluation we identified consecutive patients with RA in a clinical practice who met the first 2 ACR/EULAR criteria for remission (≤ 1 swollen joint and ≤ 1 tender joint), regardless of the value of their PtGlobal. We then analyzed the clinical characteristics associated with patients achieving the primary ACR/EULAR remission criteria (≤ 1 swollen joint, ≤ 1 tender joint, PtGlobal ≤ 1). As patients were selected because they fulfilled joint criteria, satisfying the PtGlobal criterion then became the sole determinant of remission, allowing simultaneous study of PtGlobal and ACR/EULAR remission.

MATERIALS AND METHODS

We studied consecutive patients with RA seen in a group rheumatology practice in Wichita, Kansas, USA[5], for routine rheumatology care by 1 of 4 rheumatologists between March 1, 2011, and June 6, 2011. At each examination physicians performed a 28 swollen and 28 tender joint count[6], and a physician global assessment of RA activity. The physician global assessment of RA activity question was a 0–10 numeric rating scale and was labeled "Physician's assessment of global disease activity." On that scale the following activity category groups were identified on the form in both words and numbers for the purposes of standardization: 0 = None; 1, 2, 3 = Mild; 4, 5, 6, 7 = Moderate; 8, 9, 10 = Severe. Immediately before the physician examination, each patient completed the Health Assessment Questionnaire (HAQ)-II[7] functional disability assessment, and the 21-point 0–10 visual analog scales (VAS) for pain, global severity, and fatigue. The global severity scale had the following wording, "Considering all the ways that your illness affects you, rate how you are doing on the following scale." This language is consistent with the recommendation of the ACR/EULAR criteria committee[1].

From the consecutive patients with RA examined, 77 who had ≤ 1 swollen joint and ≤ 1 tender joint were identified as eligible for ACR/EULAR remission based only on joint counts, and were studied further by a second patient questionnaire that was administered immediately after the clinic visit. The second questionnaire included a repeat VAS PtGlobal scale for the purposes of determining test-retest reliability, and an anatomic drawing modified from "The Michigan Body Map"[8] that contained 28 joints and 16 nonarticular areas that the patients could report pain in that day by checking a box. Feet, toes, ankles, hands, wrist, and fingers constituted separate, single left-side and right-side joint areas.

The next question stated: "The first question asked you about all of the ways your arthritis affects you. We want to understand what that means to you. Please check as many boxes as needed to indicate all of the factors that contributed to your severity score today." The factors that could be checked were (1) pain, (2) limitation in function, (3) fatigue, (4) problems with anxiety or depression, (5) side effects of medication, (6) cost of medical care, including medications, (7) not being able to be employed because of illness, and (8) some other factor not listed. The final question included the same 8-item checklist, but was introduced by the following text: "Now answer the same question, but check only one box below to indicate the most important factor that influenced your severity score below."

To investigate whether a PtGlobal question that asks about "all the ways your illness affects you" differs from one that specifically asks about arthritis activity, we studied a random observation from a longitudinal study of 25,051 patients with RA enrolled in the US National Data Bank for Rheumatic Diseases (NDB)[9]. We compared responses during this NDB observation to the global question described above with responses to a similar question from the Rheumatoid Arthritis Disease Activity scale[10,11] that was also included in the NDB questionnaire: "In terms of joint tenderness and swelling, how active is your arthritis today? Place an X in the box below to indicate the amount of tenderness and swelling on a scale of 0–10." Both 0–10 global scales were dichotomized at ≤ 1 and > 1 and compared by the chi-square test. The NDB has been described in detail[9].

Statistical methods. Data were analyzed by Stata, version 11.2[12]. Test-retest reliability was assessed by κ; we used the interpretation of Landis and Koch for κ values: < 0 as indicating no agreement and 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement[13]. The minimum detectable change (MDC), also known as reliable change or smallest real difference, was calculated from the PtGlobal SD and the test-retest reliability coefficient[14]. The MDC represents the smallest change in score that likely reflects true change rather than measurement error alone[14].

Exact logistic regression was used for the binary dependent variable analyses examining determinants of PtGlobal > 1 because in a number of instances, the dependent variable was completely determined by the data[15]. All data were complete (n = 77) except for physician global activity (n = 75) and erythrocyte sedimentation rate (ESR; n = 36). ESR data were missing because ESR was not collected routinely on all patients.

RESULTS

The mean age of the 77 clinical practice patients was 60.0 years (SD 14.2) and 43.4% were men. [Among all patients with RA (remission or not) seen in the clinic during the study period, the distribution by sex was men 24.9% and women 75.1%.] As expected for patients with tender and swollen joint counts ≤ 1, the VAS scores for pain [2.9 (2.1)], fatigue [3.7 (2.6)], and PtGlobal [2.7 (2.1)] were low; and the HAQ-II score was 0.71 (0.7). Remission was noted in 17 of 77 patients (22.1%) by ACR/EULAR criteria and 17.3% by AE-4 criteria; 23.5% had a physician global activity score of 0.

Table 1 compares patients who met ACR/EULAR criteria with those who did not (PtGlobal > 1). Because all of the patients had ≤ 1 swollen and ≤ 1 tender joint, meeting ACR/EULAR remission criteria depended only on having PtGlobal ≤ 1. As expected, patients with global scores > 1 had

Table 1. Comparison between American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) remission-positive and remission-negative patients who satisfy ACR/EULAR swollen and tender joint criteria (≤ 1), but differ concerning the patient global criterion (≤ 1). Data are mean (SD) unless otherwise indicated.

Variable	ACR (Global > 1) Remission (–), n = 60 (77.9%)	ACR (Global ≤ 1) Remission (+), n = 17 (22.1%)	p
Age, yrs	59.1 (12.5)	63.9 (18.8)	0.226
Sex, % male	45.8	35.3	0.443
Global severity (0–10)	3.3 (1.9)	0.4 (0.4)	< 0.001
Pain (0–10)	3.2 (1.9)	1.6 (2.3)	0.004
Fatigue (0–10)	4.3 (2.4)	1.4 (1.6)	< 0.001
HAQ-II (0–3)	0.8 (0.7)	0.3 (0.3)	0.002
MD global activity (0–10)	0.86 (0.5)	0.94 (0.8)	0.632
MD 28 tender joint count	0.25 (0.4)	0.00 (0.0)	0.021
MD 28 swollen joint count	0.13 (0.3)	0.06 (0.2)	0.089
MD global activity = 0 (%)	77.6	22.4	0.096
ACR/EULAR AE-4 remission (%)	0.0	76.5	—

HAQ-II: Health Assessment Questionnaire functional disability assessment; AE-4: ACR/EULAR-4 variable criteria.

more abnormal self-report measures, including PtGlobal, VAS pain, VAS fatigue, and HAQ-II. However, patients differed only minimally when assessed by the physician measures of physician RA global activity, and swollen and tender joint counts. Patients who met the formal ACR/EULAR remission criteria had a mean of 0 tender joints compared with 0.25 tender joints for criteria-negative patients, and this small difference reached statistical significance.

There was no agreement beyond chance ($\kappa = -0.07$, $p = 0.952$) between ACR/EULAR remission (22.1%) and physician global activity ≤ 1 (88.0%), or between ACR/EULAR remission and physician global activity of 0 (24.0%, $\kappa = 0.07$, $p = 0.276$; Table 1). To determine the test-retest reliability of the PtGlobal, we calculated the intraclass correlation coefficient (ICC) between initial and followup PtGlobal measures. The ICC was 0.84, the minimal detectable difference or reliable change was 2.3, and the standard error of the mean (SEM) was 0.84. If the ICC were 0.75, as suggested in the literature, the SEM would be 1.03. Of the 77 patients, 17 had PtGlobal ≤ 1, 27 had scores of 1.5 or 2, 7 had scores of 2.5 or 3, and 34 had scores > 1 and ≤ 3.

Correlates and predictors of PtGlobal. Because ACR/EULAR remission depended on the level of PtGlobal, we examined predictors of high PtGlobal (≥ 1) in subjects with examiner-determined tender and swollen joint counts ≤ 1 (Table 2). The number of self-reported painful joints (OR 1.49), painful nonarticular regions (OR 3.25), and the combination of both (OR 1.52) were significant predictors of high PtGlobal. Among specific sites, hand and wrist pain had an OR of 5.63 and low back pain an OR of 6.95. In addition, we noted that only 11.8% of patients in ACR/EULAR remission had hand and wrist pain compared with 43.3% not in remission. For low back pain the percentages were even more striking: 0.0% versus 23.3%.

Table 2. Effect of predictor variables on risk of high patient global score and failure to satisfy American College of Rheumatology/European League Against Rheumatism remission criteria.

Variable	OR (95% CI)	p
Painful joints and nonarticular regions (0–44)	1.49 (1.10, 2.28)	0.005
Painful joints (0–28)	1.49 (1.05, 2.30)	0.018
Painful nonarticular regions (0–16)	3.25 (1.03, 27.30)	0.040
Specific painful regions		
Hands and wrists	5.63 (1.15, 55.06)	0.028
Low back	6.95 (1.06, ∞)	0.041
Shoulders	4.36 (0.57, 199.65)	0.250
Elbows	2.00 (0.26, ∞)	0.553
Hip	1.55 (0.18, ∞)	0.721
Upper back	1.55 (0.18, ∞)	0.721
Knees	1.39 (0.36, 6.64)	0.848
Feet	1.28 (0.33, 6.16)	0.955
Buttocks	0.69 (0.05, ∞)	1.000
Groin	1.77 (0.19, 86.81)	1.000
Jaw	0.28 (0.007, ∞)	1.000
Neck	1.45 (0.15, 73.04)	1.000
ESR ≥ 20 (male), ≥ 30 (female)	0.40 (0.03, ∞)	1.000*
VAS fatigue (0–10)	2.19 (1.45, 3.77)	0.000
VAS pain (0–10)	1.68 (1.18, 2.58)	0.002
HAQ-II disability (0–3)	9.33 (2.22, 54.42)	0.001

* n = 33. ESR: erythrocyte sedimentation rate; VAS: visual analog scale; HAQ-II: Health Assessment Questionnaire functional disability assessment.

Among the 17 patients who had pain or swelling on the physician hand or wrist examination, 8 (47.1%) had hand or wrist pain by patient self-report. Among the 56 with a negative physician hand and wrist examination, 24 (30.4%) had hand or wrist pain by self-report. There was no statistical association between physician and patient positive hand and wrist examinations ($p = 0.204$).

We also found that VAS fatigue (OR 2.19), VAS pain (OR 1.68), and HAQ-II (OR 9.33) were significant predictors of high PtGlobal (Table 2). Although ESR data were available from only 33 patients, only 2 patients had high ESR levels. Thus 100% of 5 patients in ACR/EULAR remission had normal ESR levels and 92.9% of 28 patients not in remission had normal ESR levels.

Further insight into the importance of patient self-report factors can be seen in Table 3, which summarizes the associations of the primary self-report issues determining high PtGlobal. Pain, functional limitations, and fatigue were among the most important determinants of PtGlobal, present in 52.0% to 59.7% of respondents. But pain and fatigue were not contributing factors for 40.3% and functional limitations were not contributing factors for 48.1% of patients. Pain was the most important factor for 38.2%, functional limitations for 25.0%, and fatigue for 26.5%.

Figure 1 shows the relationship between PtGlobal and

Table 3. Primary and contributing factors influencing the patient severity score.

Factor	Most Important Factor, %	Contributing Factor, %	Not a Contributing Factor, %
Pain	38.2	59.7	40.3
Functional limitations	25.0	52.0	48.1
Fatigue	26.5	59.7	40.3
Mood	2.9	14.3	85.7
Medication side effects	1.5	15.6	84.4
Cost of medical care	2.9	11.7	88.3
Work disability	1.5	11.7	88.3
Other	1.5	3.9	96.1

fatigue, pain, and HAQ-II, according to whether physicians rated RA activity (physician global) ≤ 1 or > 1. The vertical line at 1.0 divides patients who meet ACR/EULAR criteria from those who do not. As noted, there is minimal connection between high physician global activity (> 1) and high PtGlobal (> 1); fatigue (correlation r = 0.773), pain (r = 0.770), and HAQ-II (r = 0.736) are strongly related to PtGlobal.

Many of the patients who failed to meet ACR/EULAR remission criteria appeared to do so because of non-RA activity-related problems. Of the 18 patients who had a physician global activity score of 0 and who had no swollen or tender joints and who were, therefore, almost certainly in clinical remission, 13, or 72.2%, did not satisfy ACR/EULAR remission criteria because of PtGlobal (Figure 2, right). Of the 66 patients who had ≤ 1 swollen joints, ≤ 1 tender joints, and a physician global score ≤ 1, 79.3% did not meet ACR/EULAR remission criteria (Figure 2, right panel), because PtGlobal was > 1. When this group was examined further by including patients with ESR results (≤ 1 swollen joints, ≤ 1 tender joints, physician global ≤ 1, and ESR < 20 mm/h in men and < 30 mm/h in women), 82.8% of 29 patients did not meet ACR/EULAR criteria (Figure 2, left panel). The exclusive distribution of patients with low back pain to high PtGlobal can be seen in Figure 2, right panel. Also of note is a subset of patients with high PtGlobal, pain, fatigue, and HAQ-II at levels of physician global activity ≤ 1 (Figure 1).

Questionnaire wording. It seemed possible that the PtGlobal question used in most clinical trials and recommended by the ACR/EULAR committee measured RA severity and activity rather than just RA activity, and that a rewording might improve the functioning of the global question as a measure of remission. To investigate whether a "severity scale" led to more abnormal scores than an "activity scale," we obtained data from 25,051 patients with RA followed in the NDB and compared the severity and activity questions after dichotomizing both scales at ≤ 1 and > 1. The percentage of patients > 1 was 76.5% for the activity scale and 79.8% for the severity scale (p < 0.001). This suggests that the use of a severity scale slightly increases (by 3.3%) the probability of an abnormal (> 1) PtGlobal.

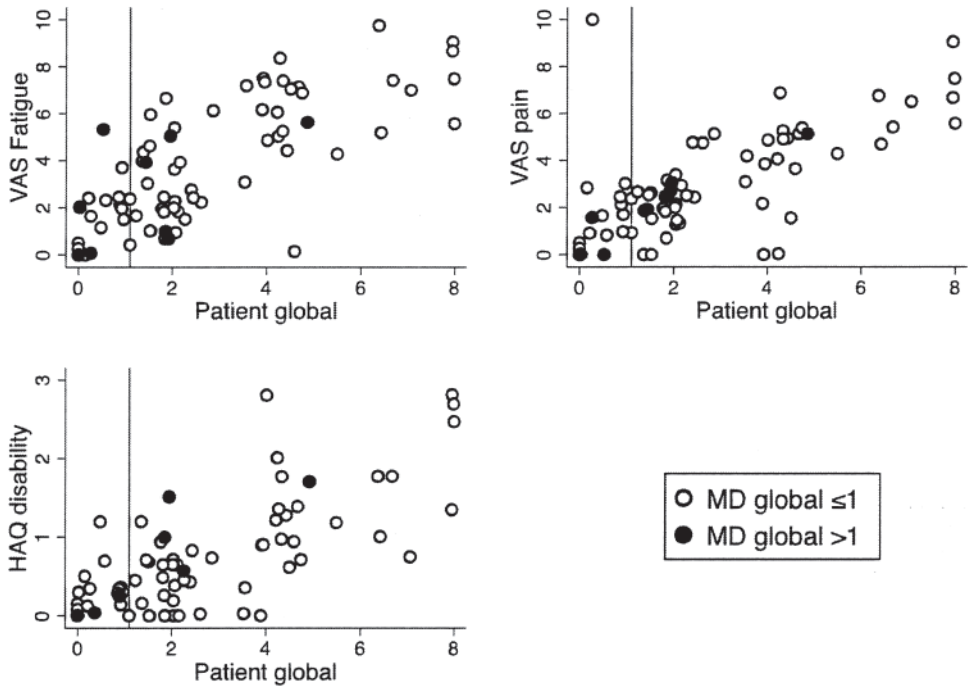

Figure 1. The relation between patient global health assessment and fatigue (left) and pain (right) according to physician (MD) global activity status. The vertical line separates patients in ACR/EULAR remission (left of line) from those not in remission. A small amount of random noise is added to the data points to improve visibility by minimizing overlap. VAS: visual analog scale; HAQ: Health Assessment Questionnaire.

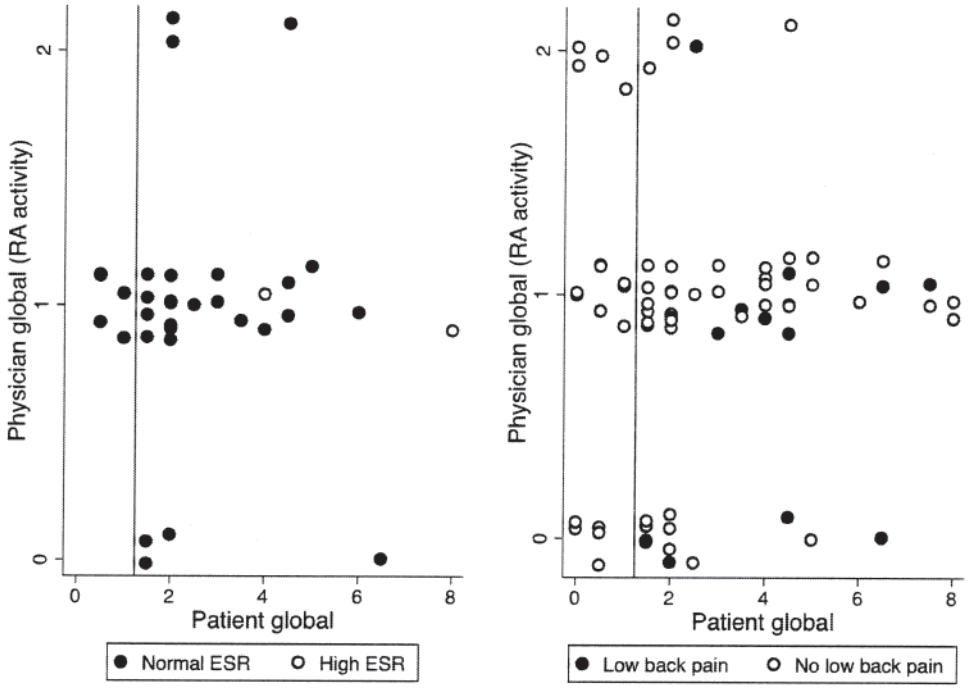

Figure 2. The relation between the levels of physician global activity and patient global health assessment for all patients (right panel) and for the subset of patients for whom erythrocyte sedimentation rate (ESR) was available (left panel). A small amount of vertical random noise is added to the physician global activity points to improve visibility by minimizing overlap. The vertical line separates patients in ACR/EULAR remission (left of line) from those not in remission. RA: rheumatoid arthritis.

DISCUSSION

Although there have been many publications that addressed the issue of remission in RA[1,5,16,17,18,19,20,21,22], the ACR/EULAR criteria were designed to be rigorous, endorsed criteria that resolved the controversies inherent in previous definitions. Faced with the difficult task of formulating remission criteria based on clinical findings, the ACR/EULAR committee studied data from randomized clinical trials, as well as survey data on threshold levels for remission obtained from 27 committee members (25 experienced RA clinical researchers and 2 patients)[1]. PtGlobal was added to the remission definition because it performed well in the reviewed clinical treatment trials. The criteria authors noted that "PtGlobal assessment and patient-reported pain were statistically significant predictors that discriminated between treatments after controlling for physician-reported measures (tender joint count and swollen joint count) and a laboratory measure (C-reactive protein)," and that global "was the best predictor of treatment assignment among all outcomes in one trial and the fourth best of core set measures[23] in another." In addition, PtGlobal appeared to "mitigate the limitation of using a 28-joint count" and "prevent misclassification." Of the 27 committee members surveyed as to the highest level of PtGlobal that would be compatible with remission if all other measures suggested remission (Felson, *et al*[1]; Table 1), the 50th percentile of responses was 2 and the 80th percentile was 3. If PtGlobal were the only measure assessed, the surveyed committee members' value for global was 1 at the 50th and 80th percentiles. The actual value selected by the remission committee for incorporation into the full criteria was 1.

In our current study of consecutive patients with RA in clinical practice who met the first 2 ACR/EULAR criteria for remission (swollen joints ≤ 1 and tender joints ≤ 1), 17 (22.1%) also had a PtGlobal ≤ 1 and therefore met full ACR/EULAR criteria for remission in clinical practice. A primary objective of our study was to examine the 60 (77.9%) that had a high PtGlobal (> 1) and did not meet ACR/EULAR criteria even though they met ACR/EULAR remission joint criteria. Although we did not evaluate the AE-4 criteria in this report, the PtGlobal would have been exactly the same if we had studied this criteria definition in detail because all patients not satisfying ACR/EULAR criteria would also not satisfy AE-4 criteria. We noted that 22.1% of patients satisfied ACR/EULAR criteria and of those, 76% would also satisfy AE-4 criteria. Thus these data show that the probability of satisfying AE-4 criteria can never be greater than the probability of satisfying ACR/EULAR criteria, and most often will be lower.

An important finding of this study was that many of the patients who failed to meet ACR/EULAR remission criteria appeared to do so because of non-RA activity-related prob-

lems. Evidence for the role of noninflammatory problems comes most clearly from 18 patients who had a physician global of 0 and who had no swollen or tender joints and who were, therefore, almost certainly in clinical remission, as well as from other groups with physician global of 0 or 1 (Figure 2). A normal ESR (ACR/EULAR committee definition) was observed in 94.3% of patients. Low back pain was noted in 18.2% of patients, all of whom had PtGlobal > 1. Low back pain may be an important factor in reporting severity in RA, as it has been identified in 19.9% to 34.4% of patients in other studies[24,25,26], depending on the setting and definitions; it is always associated with worse self-reported clinical variables. As suggested by the results of Table 2 and the patient preferences of Table 3 and data of Figure 1, functional status and fatigue played an important role in PtGlobal levels. Although increased in RA, fatigue is a highly prevalent symptom in noninflammatory conditions[27] and is associated with increased self-assessment severity. Similarly, functional status can be abnormal in the presence of joint damage in settings when inflammation is not present[28].

Could some patients with high global scores have fibromyalgia? Although we have insufficient data to make this determination accurately, we might suspect this in patients with levels of fatigue ≥ 6[29]. Figure 1 identifies many patients with high fatigue. Whether this is fibromyalgia or not, Figure 1 shows many patients with high levels of pain, fatigue, and HAQ who have PtGlobal > 1.

Figures 1 and 2 provide information on one other important area: the question of the reliability of PtGlobal by showing a large cluster of patients with scores > 1 and ≤ 3. In our previous longitudinal report on ACR/EULAR remission[3] we noted that some patients seemed to oscillate between meeting the global criterion and not meeting it. In our current study, we found the PtGlobal reliability coefficient to be 0.84, although other studies have shown it to be 0.75[30], perhaps reflecting the relatively short time interval for the test-retest used in this study. Reliability coefficients ≥ 0.90 are usually considered to be necessary for use in individual patients[14]. We also determined that the minimal detectable change and SEM for PtGlobal was 2.3 and 0.8 at the observed ICC of 0.84. At an ICC of 0.75[30], the MDD and SEM would be 2.9 and 1.03, respectively. These data suggest the possibility that some to many of the clustered values around a PtGlobal of 1.5 to 2, noted best in Figures 1 and 2, might really represent "true" scores of ≤ 1 and that the "true" number of patients in ACR/EULAR remission might be greater than observed. It is of interest that if the limit of PtGlobal for remission were raised from 1 to 2, as discussed in the ACR/EULAR remission report, the number of patients in remission in our current study would more than double, from 22.1% to 57.1%.

Although we found that a PtGlobal activity scale was 3.3% less likely to be abnormal (> 1) compared with a global severity scale, this difference is small and could not have affected the study results in a meaningful way. In addition, the use of the current wording makes it possible to use the PtGlobal question in illnesses where activity is not part of the illness (e.g., osteoarthritis, fibromyalgia). In addition, a 2011 study that evaluated 5 forms of PtGlobal/activity found little clinical difference among the scales[31].

Overall, the data of our study show that it is probable that the ACR/EULAR remission criteria when used in the clinic underestimate the true rate of remission as contemplated by the ACR/EULAR committee. This occurs because of (1) the exclusion of patients with elevated global scores not caused by RA activity, and (2) measurement error. We would estimate that at least 50% of ACR/EULAR criteria-negative patients in our study are probably truly in remission.

There are a number of limitations to our report. There is no gold standard definition of remission against which the ACR/EULAR definition and our data can be compared. Consequently, we cannot make absolute judgments regarding which patients were or were not in remission. Even a physician global of 0, which meant "no activity" and was written as such on the examination form, is not universally accepted. However, a physician global of 0 together with no swollen or tender joints has a high probability of identifying patients in remission. It was also difficult for us to draw inferences about patient-reported hand and wrist pain because for reasons of feasibility we lumped together all the joints of the hand and wrist on the patient self-assessment form. In addition, while we used 8 domains of importance for factors influencing PtGlobal (Table 3), other and different domains have been suggested[32,33], and it is possible the results of Table 3 might have been slightly different if we had incorporated those domains.

In our previous report on the use of remission criteria in clinical practice[3], we demonstrated problems with the reliability of the physician joint examination, noting that some physician joint scores could vary by a factor of 2, depending on the examiner. The current report illuminates the degree to which PtGlobal at the level of the individual patient is limited by reliability issues. In addition, false-negative remission appears to be a major problem with the ACR/EULAR remission criteria. It seems possible that increasing the PtGlobal threshold from ≤ 1 to ≤ 2 might more accurately identify remission, but such a change might not be in line with the concept of remission put forth by the ACR/EULAR committee. We suggest, therefore, that in clinical practice any of the ACR/EULAR remission criteria for clinical practice be used as rough guides, perhaps as measures of low disease activity rather than an absolute state, particularly if remission status is to be reviewed by regulatory or third-party payers.

We would emphasize that many of the problems we noted in clinical practice may not apply to the randomized trial setting, where problems of reliability are understood and overcome by large sample sizes. In addition, it seems likely that PtGlobal has a different meaning to patients in clinical trials compared with clinical practice: patients in trials almost always begin with high levels of RA activity and changes in

PtGlobal most likely refer to the change in activity. By contrast, in clinical practice the memory of greater activity, which may have occurred years before, probably does not play a similar role. Rather, nonarticular pain, function, and fatigue may be more important among those already satisfying the ACR/EULAR joint criteria.

REFERENCES

1. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum 2011;63:573-86.
2. Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. Best Pract Res Clin Rheumatol 2007;21:663-75.
3. Shahouri SH, Michaud K, Mikuls TR, Caplan L, Shaver TS, Anderson JD, et al. Remission of rheumatoid arthritis in clinical practice: Application of the American College of Rheumatology/European League Against Rheumatism 2011 remission criteria. Arthritis Rheum 2011;63:3204-15.
4. Khan NA, Spencer HJ, Abda E, Aggarwal A, Alten R, Ancuta C, et al. Determinants of discordance in patients' and physicians' rating of rheumatoid arthritis disease activity. Arthritis Care Res 2012;64:206-14.
5. Shaver TS, Anderson JD, Weidensaul DN, Shahouri SS, Busch RE, Mikuls TR, et al. The problem of rheumatoid arthritis disease activity and remission in clinical practice. J Rheumatol 2008;35:1015-22.
6. Fuchs HA, Brooks RH, Callahan LF, Pincus T. A simplified twenty-eight-joint quantitative articular index in rheumatoid arthritis. Arthritis Rheum 1989;32:531-7.
7. Wolfe F, Michaud K, Pincus T. Development and validation of the Health Assessment Questionnaire II: A revised version of the Health Assessment Questionnaire. Arthritis Rheum 2004;50:3296-305.
8. Brummett CM, Hassett AL, Brummett KA, Clauw DJ, Williams DA. The Michigan Body Map and its use in assessing the American College of Rheumatology survey criteria for fibromyalgia. Arthritis Rheum 2012; (in press).
9. Wolfe F, Michaud K. The National Data Bank for rheumatic diseases: A multi-registry rheumatic disease data bank. Rheumatology 2011;50:16-24.
10. Fransen J, Langenegger T, Michel BA, Stucki G. Feasibility and validity of the RADAI, a self-administered rheumatoid arthritis disease activity index. Rheumatology 2000;39:321-7.
11. Stucki G, Liang MH, Stucki S, Bruhlmann P, Michel BA. A self-administered rheumatoid arthritis disease activity index (RADAI) for epidemiologic research. Psychometric properties and correlation with parameters of disease activity. Arthritis Rheum 1995;38:795-8.
12. Stata C. Stata Statistical Software: Release 11.2. College Station, TX, USA: Stata Corporation; 2009.
13. Landis J, Koch G. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
14. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. J Clin Epidemiol 2004;57:1008-18.
15. Mehta CR, Patel NR. Exact logistic regression: Theory and examples. Stat Med 1995;14:2143-60.
16. Pemberton R, Peirce EG. A clinical and statistical study of chronic arthritis based on eleven hundred cases. Am J Med Sci 1927;173:31-58.
17. Thompson HE, Wyatt BL, Hicks RA. Chronic atrophic arthritis. Ann Intern Med 1938;11:1792-805.
18. Pinals RS, Baum J, Bland J, Fosdick WM, Kaplan SB, Masi AT, et al. Preliminary criteria for clinical remission in rheumatoid arthritis. Arthritis Rheum 1981;24:1308-15.
19. Sokka T, Hetland ML, Makinen H, Kautiainen H, Horslev-Petersen K, Luukkainen RK, et al. Remission and rheumatoid arthritis: Data on patients receiving usual care in twenty-four countries. Arthritis Rheum 2008;58:2642-51.
20. van Tuyl LH, Vlad SC, Felson DT, Wells G, Boers M. Defining remission in rheumatoid arthritis: Results of an initial American College of Rheumatology/European League Against Rheumatism consensus conference. Arthritis Rheum 2009;61:704-10.
21. Aletaha D, Ward MM, Machold KP, Nell VP, Stamm T, Smolen JS. Remission and active disease in rheumatoid arthritis: Defining criteria for disease activity states. Arthritis Rheum 2005;52:2625-36.
22. Mierau M, Schoels M, Gonda G, Fuchs J, Aletaha D, Smolen JS. Assessing remission in clinical practice. Rheumatology 2007;46:975-9.
23. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis Rheum 1993;36:729-40.
24. Hagen KB, Kvien TK, Bjørndal A. Musculoskeletal pain and quality of life in patients with noninflammatory joint pain compared to rheumatoid arthritis: A population survey. J Rheumatol 1997;24:1703-9.
25. Neva MH, Häkkinen A, Isomäki P, Sokka T. Chronic back pain in patients with rheumatoid arthritis and in a control population: Prevalence and disability — a 5-year follow-up. Rheumatology 2011;50:1635-9.
26. Wolfe F. Determinants of WOMAC function, pain and stiffness scores: Evidence for the role of low back pain, symptom counts, fatigue and depression in osteoarthritis, rheumatoid arthritis and fibromyalgia. Rheumatology 1999;38:355-61.
27. Bergman MJ, Shahouri SS, Shaver TS, Anderson JD, Weidensaul DN, Busch RE, et al. Is fatigue an inflammatory variable in rheumatoid arthritis (RA)? Analyses of fatigue, osteoarthritis, and fibromyalgia. J Rheumatol 2009;36:2788-94.
28. Bruce B, Fries J. Longitudinal comparison of the Health Assessment Questionnaire (HAQ) and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). Arthritis Rheum 2004;51:730-7.
29. Katz RS, Wolfe F, Michaud K. Fibromyalgia diagnosis: A comparison of clinical, survey, and American College of Rheumatology criteria. Arthritis Rheum 2006;54:169-76.
30. Lassere MN, van der Heijde D, Johnson KR, Boers M, Edmonds J. Reliability of measures of disease activity and disease damage in rheumatoid arthritis: Implications for smallest detectable difference, minimal clinically important difference, and analysis of treatment effects in randomized controlled trials. J Rheumatol 2001;28:892-903.
31. Dougados M, Ripert M, Hilliquin P, Fardellone P, Brocq O, Brault Y, et al. The influence of the definition of patient global assessment in assessment of disease activity according to the Disease Activity Score (DAS28) in rheumatoid arthritis. J Rheumatol 2011;38:2326-8.
32. Gossec L, Dougados M, Rincheval N, Balanescu A, Boumpas DT, Canadelo S, et al. Elaboration of the preliminary Rheumatoid Arthritis Impact of Disease (RAID) score: A EULAR initiative. Ann Rheum Dis 2009;68:1680-5.
33. Sanderson T, Morris M, Calnan M, Richards P, Hewlett S. Patient perspective of measuring treatment efficacy: The rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. Arthritis Care Res 2010;62:647-56.