Short-term Outcome After Anti-Tumor Necrosis Factor-α Therapy in Rheumatoid Arthritis: Do We Need to Revise Our Assessment Criteria?

YASSER EL MIEDANY, SALLY S. YOUSSEF, and MAHA EL GAAFARY

ABSTRACT. ***Objective.*** To evaluate the Disease Activity Score (DAS) using various aggregated dimensions to quantify treatment outcome in patients with rheumatoid arthritis (RA), in order to determine the best instrument to be used as an endpoint that indicates good response in terms of EULAR response criteria and DAS28 remission criteria, and which satisfies the demands of clinical rheumatology.

Methods. Using raw data for each patient subjected to anti-tumor necrosis factor-α therapy (81 patients), before and 6 months after treatment, DAS28 was calculated 4 times using the standard equation, as follows: (1) DAS 1 (the standard DAS28): tender joint count (TJC), swollen joint count (SJC), patient global assessment (PGA), erythrocyte sedimentation rate (ESR); (2) DAS 2: TJC, SJC, PGA, C-reactive protein (CRP); (3) DAS 3: TJC, SJC, physician global assessment (PhGA), ESR; and (4) DAS 4: TJC, SJC, PhGA, CRP. Disease activity was identified if DAS score exceeded 5.1. A clinically significant response was recorded if there had been improvement of > 1.2 of the DAS score.

Results. DAS 2, DAS3, and DAS4 were superior to the current DAS score used for assessment of RA activity (effect size differences were –0.35, –0.13, and –0.48, respectively). Assessment of disease activity using TJC, SJC, PhGA, and CRP was the best tool to assess response to therapy. ESR was marginally superior to CRP in its sensitivity to monitor disease activity changes (effect sizes 1.08 and 1.03, respectively).

Conclusion. These results suggest that self-report indices on their own, such as PGA and pain score, are inadequate indicators of disease activity. The DAS might profitably be amended by one or 2 continuous measures for better quantification of the degree of improvement of patients on a given therapeutic modality. Using PhGA and CRP instead of PGA and ESR, respectively, in the DAS equation discriminated better between different patients' responses than the traditional DAS score. (J Rheumatol 2006;33:490–6)

Key Indexing Terms:
DISEASE ACTIVITY SCORE RHEUMATOID ARTHRITIS BIOLOGIC THERAPY

Outcome measures for rheumatic diseases have been a major focus of clinical research for decades. A single measure was not available to serve as a gold standard to assess clinical status in patients with rheumatoid arthritis (RA). The original measures for disease activity and severity for RA such as the Lansbury[1] and Ritchie[2] activity indices and the American Rheumatism Association functional classification have been modified to generate scales that satisfy the demands of clinical epidemiology. The American College of Rheumatology (ACR) and the European League of Associations for Rheumatology (EULAR) developed a newer activity index[3] that incorporated different dimensions, each of which has been subject to scrutiny for both validity and reliability. A widely used index in studies of patients with RA is the Disease Activity Score (DAS)[4-6]. The DAS takes into consideration the tender joint count (TJC), the swollen joint count (SJC), erythrocyte sedimentation rate (ESR), and patient global assessment of general health (PGA). The DAS determines whether a patient has had a clinically significant response to therapy (i.e., a dichotomous decision) based on achieving 1.2 improvement of the patient score[6]. While the DAS has undergone major efforts to validate it as a discriminatory and specific outcome measure for RA, questions remain about its utility and appropriateness, especially for the newer therapeutic agents such as tumor necrosis factor-α (TNF-α) inhibitors. A primary concern is that, apart from the TJC and SJC, a discrepancy has been noted on recording the variables assessed by visual analog scales (VAS) such as the pain score, the PGA, and the physician global assessment (PhGA) as well as measures of acute phase reactants such as ESR and C-reactive

From the Rheumatology and Rehabilitation Department, and the Department of Community, Environmental and Occupational Medicine, Faculty of Medicine, Ain Shams University, Cairo, Egypt.

Y. El Miedany, MD, Professor, Rheumatology and Rehabilitation; S.S. Youssef, MD, Assistant Lecturer, Rheumatology and Rehabilitation; M. El Gaafary, MD, Lecturer, Community, Environmental and Occupational Medicine, Faculty of Medicine.

Address reprint requests to Dr. Y. El-Miedany, 6 Blenheim Close, Meopham, Kent DA13 0PQ, England.
E-mail: yasser_elmiedany@yahoo.com

Accepted for publication October 18, 2005.

protein (CRP); this raises the question whether the DAS has limited utility in discriminating among efficacious agents. This observation has significant clinical implications, because based on the DAS score the physician will determine the response to drug therapy and the possibility of modifying the dose or frequency of injection, as well as the possibility for adding another disease modifying drug (DMARD) therapy to the current biologic therapy. Finally, there is no assurance that the scale behaves in a linear fashion, so that the current trend of using it as a more stringent test of efficacy may not result in appropriate statistical test characteristics[7].

With these questions in mind, we undertook a reevaluation of the DAS examining various aggregated forms. Our intent was to determine which of the various forms of DAS would provide the best assessment of the short-term outcome of anti-TNF therapy, that is, the best tool to be used as an endpoint that would show good response by EULAR criteria and DAS28 remission, and also satisfy the demands of clinical rheumatology. This could be achieved by assessing the effect size and the relative sensitivity to change of the different measures.

MATERIALS AND METHODS

Patients. We used raw data for each patient with RA subjected to anti-TNF therapy, before and 6 months after treatment; there were 81 patients (71 women, 10 men) diagnosed according to the ACR criteria of 1987[8]. Biologic therapy was started after failure of 2 DMARD, one of them methotrexate (MTX). The anti-TNF therapy used was infliximab (39 patients) and etanercept (42 patients).

Local ethical and methodological protocols for approval of the study were followed.

Disease activity assessment. The DAS was calculated 4 times using various aggregated indicators as follows. (1) The first indicator measured improvement in TJC by scoring tenderness to pressure and joint manipulation on physical examination; the types of tenderness were combined into a single tender versus nontender dichotomy for each point. The scores for each patient were summed over 28 joints. (2) The second indicator measured improvement in SJC. Analogous to the TJC, the scores of 28 joints for each patient were summed. (3) The third indicator measured the patient's global health assessment of disease activity on a continuous 0–100 mm VAS. (4) The fourth indicator was the ESR measured by the Westergren method. In addition to these standard measures, other indicators of disease activity were also recorded, including patient assessment of pain and the PhGA of disease activity using a continuous 0–100 mm VAS. The level of CRP (mg/l) by ELISA was also recorded.

These 6 indicators were recorded twice: once at baseline and once after 6 months of therapy. We examined various ways of using the scores that would best reflect the ability to detect differences between baseline and after 6 months of anti-TNF therapy. The first 2 indicators (TJC and SJC) were standard in all measures assessed. The first outcome measure tested (DAS1) was calculated using the standard DAS equation:

$$DAS = 0.56\ (\sqrt{28TJC}) + 0.28\ (28SJC) + 0.70\ (\text{Log n ESR}) + 0.014\ (GH)$$

where GH = patient global health. This definition uses 4 measures that are qualified with a baseline score that can be compared with a final score taken after 6 months of therapy. We studied 3 other ways in calculation of the DAS score using the ReDAS software program[9], before and 6 months after biologic therapy, as follows.

DAS2: TJC, SJC, CRP, and PGA.

$$DAS2 = 0.56\ (\sqrt{28TJC}) + 0.28\ (28SJC) + 0.70\ (\text{Log n CRP}) + 0.014\ (PGA)$$

DAS3: TJC, SJC, ESR, and PhGA.

$$DAS3 = 0.56\ (\sqrt{28TJC}) + 0.28\ (28SJC) + 0.70\ (\text{Log n ESR}) + 0.014\ (PhGA)$$

And DAS4: TJC, SJC, CRP, and PhGA.

$$DAS4 = 0.56\ (\sqrt{28TJC}) + 0.28\ (28SJC) + 0.70\ (\text{Log n CRP}) + 0.014\ (PhGA)$$

Disease activity was identified if DAS score exceeded 5.1. A clinically significant response was recorded if there had been improvement of > 1.2 of the DAS score. Sensitivity to change or responsiveness of the 4 DAS scores was tested to assess how far each test was able to detect changes in clinical status when they occurred. Tests that are sensitive to change will register large changes when a patient's clinical status improves, while insensitive measures mean relatively stable results despite noticeable clinical improvement. Using these 6 indicators, we also compared the criteria that achieved thresholds of 1.8 and 2.4 improvement of the DAS.

Statistical analysis. The effect size was used as the measure of sensitivity to change. It was computed as (baseline value – followup value)/baseline standard deviation. Positive effect sizes indicate that the measured variable decreases with treatment, negative effect sizes indicate that the measured variable increases with treatment, and effect sizes of zero indicate that treatment has no effect on the variable. Effect size was calculated for each suggested DAS score in relation to the standard DAS.

To compare the sensitivity to change between different measured variables, the effect size of the new sets of indicators (DAS2, DAS3, and DAS4) was subtracted from the effect size of the standard one (DAS1) to yield an effect size difference. Positive effect size difference showed that the standard set of indicators is more sensitive than the new ones, while a negative effect size difference showed that the new indicators were more sensitive to change than the standard one. For CRP and ESR, the effect size of CRP was subtracted from the effect size of ESR. Positive effect size difference indicated that ESR is more sensitive than CRP, while a negative effect size difference indicated that CRP was more sensitive to change than ESR. Ninety-five percent confidence intervals (95% CI) of the effect size difference that excluded zero indicated a significant difference (at a type I error rate of 0.05) in the effect size.

All statistical analysis was performed using the 11th version of the SPSS statistical package. Paired t test, chi-square, and Pearson correlation coefficients were used to test all the presented associations and correlations. Fisher's exact test was used if fewer than 5 observations were expected in any cell of the 2 by 2 tables. Kappa statistic was used to evaluate the degree of agreement between the different variants of the DAS at different cutoff points.

RESULTS

This study included 81 patients with RA. Their mean age was 58.6 (SD 6.6) years (range 43–76). Seventy-one patients (87.7%) were women. Their average duration of illness was 13.5 years (range 4–26).

Comparing the baseline record of each of the DAS score variants with the records 6 months after the start of anti-TNF therapy, there was a significant improvement of tender and swollen joint counts, associated with similar significant improvement of all the other criteria denoting positive response to biologic therapy (Table 1). A total of 59 (72.8%) patients showed a DAS exceeding the cutoff of 1.2.

Table 2 shows the mean values of DAS2, DAS3, and DAS4 in relation to the groups of the original DAS that were improved and not improved. DAS3 presented the lowest lower limit for improvement (2.06), while DAS4 presented

Table 1. Effect of 6 months of biologic therapy on different disease measures.

Measure	Before Treatment	After 6 months	p
TJC	10.9 ± 5.0	5.7 ± 3.9	< 0.001
SJC	3.9 ± 2.2	0.91 ± 1.3	< 0.001
PGA	72.7 ± 18.5	58.0 ± 23.8	< 0.01
Pain score	74.7 ± 16.9	57.7 ± 23.8	< 0.01
PhGA	64.9 ± 18.9	27.6 ± 13.0	< 0.001
ESR	52.7 ± 19.4	31.8 ± 14.2	< 0.001
CRP	48.1 ± 30.7	16.5 ± 13.4	< 0.01
DAS1	6.1 ± 0.9	4.5 ± 1.0	< 0.001
DAS2	5.9 ± 1.0	3.8 ± 1.4	< 0.001
DAS3	5.9 ± 1.0	4.0 ± 0.9	< 0.001
DAS4	5.8 ± 1.1	3.4 ± 1.3	< 0.001

TJC: tender joint count, SJC: swollen joint count, PGA: patient global health assessment, PhGA: physician global health assessment, ESR: erythrocyte sedimentation rate, CRP: C-reactive protein.

Table 2. Average DAS2, DAS3 and DAS4 scores in relation to the cutoff of improvement of DAS1.

	< 1.2 "Not improved," mean (95% CI)	≥ 1.2 "Improved," mean (95% CI)
DAS2	0.86 (0.47–1.25)	2.55 (2.38–2.72)
DAS3	0.96 (0.68–1.24)	2.20 (2.06–2.34)
DAS4	1.23 (0.72–1.74)	2.89 (2.76–3.02)

the highest lower limit (2.76). However, identification of a cutoff point for declaration of improvement may need further analysis, which was beyond the main concern of this study.

Analysis was also carried out to study the correlation of the different indicators with each other before and after 6 months of drug therapy. Table 3 shows that prior to therapy there was a significant correlation between TJC and (1) the physician global health assessment, (2) pain score, (3) the patient global health assessment, and (4) SJC, as well as both CRP and ESR (in order of importance). Also, there was a significant correlation between the PhGA and both PGA and pain score. PhGA was perfectly correlated with CRP, SJC, ESR, pain score, and PGA.

These data permitted further analysis of the effect size of each of the suggested DAS scores (DAS2, DAS3, DAS4), and the effect size difference was calculated in relation to the traditional DAS (DAS1) (Tables 4 and 5). All the suggested DAS scores (DAS2, DAS3, and DAS4) showed their superiority over the traditional DAS1, as the effect size differences were all negative, and the differences were all significantly higher than the standard DAS. DAS3 reported the smallest effect size difference.

Studying the effect size as well as the effect size difference for both ESR and CRP (Table 6) showed that there was no significant difference in terms of sensitivity to monitor disease activity changes between the 2 tests, although there was marginal nonsignificant superiority of ESR over CRP.

In all variants of the DAS scores, duration of illness was negatively correlated with the effect size of treatment, meaning that a long history of disease was associated with a minimal effect of therapy (Table 7). The same finding was noted with ESR. In contrast, CRP did not show significant correlation with the duration of illness, either before or after treatment. Further, age did not show any significant association with the size of treatment effect.

DISCUSSION

Management of patients with RA is critically dependent on the use of appropriate outcome measures. These measures need to fulfil various functions. First, they should be specific in detection of clinical response with potential for therapeutic utility. In addition, they should be able to discriminate between responders and nonresponders to generate at least a rank order of effectiveness or to enable modification of the therapeutic regimen. Finally, the measures should be responsive to modest but clinically important differences and they should be interpreted in terms of clinical benefit[10]. There were some points of concern about the design of this study: first, the use of parameters similar to the ones already used in the traditional DAS equation. Second, considering the contradiction between functional disability [as assessed by the Health Assessment Questionnaire (HAQ)] and disease activity as measured by the traditional DAS score[10], we were left with 3 variables: pain score, physician global assessment (both measured on a VAS), and CRP. However, the main basis for this study was the contradiction in assessment of the different disease activity measures, and this raised the question, do we calculate them in the wrong way?

In the 1980s there was a concentrated effort to determine what outcome measures should be used for RA clinical trials. The introduction of patient-based outcome measures, including patient global assessment and pain scores[11], led to a core set of disease activity measures for RA. However, in contrast with the ACR scoring system, the DAS does not include the pain score, physician global assessment, HAQ, or CRP level. Furthermore, the traditional DAS was calculated based on studies carried out on RA patients treated with DMARD. DMARD in those studies were somewhat less effective than the agents currently used. This resulted in a more stringent testing of the outcome measures than might be the case with more effective agents. Our study was not merely an attempt to evaluate the efficacy of the drugs being administered; our aim was to determine whether there is a more effective way of using the same information to maximize the ability to discriminate between the different outcomes of the same therapy and to determine the minimal degree of improvement. Thus the exact extent of the data set is secondary to our ability to use these data to evaluate the various scoring systems.

Initial analysis of our results showed that the favorable parameters of the quantitative measures to assess disease activity and monitor drug therapy were the tender and swollen

Table 3. Correlation between the disease activity measures assessed before and after 6 months of biologic therapy.

	CRP	ESR	PhGA	Pain Score	PGA	SJC
TJC						
Baseline	0.403**	0.303**	0.709**	0.623**	0.622**	0.598**
After 6 mo	–0.05	0.321**	0.364**	0.141	0.133	0.353*
SJC						
Baseline	0.356**	0.378**	0.560**	0.421**	0.415**	
After 6 mo	0.560**	0.394**	0.687**	0.304*	0.308*	
PGA						
Baseline	–0.075	0.084	0.761**	0.975**		
After 6 mo	0.473**	0.562**	0.601**	0.997**		
Pain score						
Baseline	–0.082	0.066	0.734**			
After 6 mo	0.472**	0.557**	0.604**			
PhGA						
Baseline	0.445**	0.480**				
After 6 mo	0.831**	0.643**				
ESR						
Baseline	0.732**					
After 6 mo	0.844**					

* $p < 0.05$. ** $p < 0.001$. Baseline measures were correlated with their baseline and the 6 mo measures with their corresponding 6 month measures. TJC: tender joint count, SJC: swollen joint count, PGA: patient global health assessment, PhGA: physician global health assessment, ESR: erythrocyte sedimentation rate, CRP: C-reactive protein.

Table 4. Correlation of different DAS effect sizes (ES) with disease measures at baseline and after 6 months of therapy. DAS3 effect size shows the best correlation with the measured variables at baseline, while DAS2 effect size had the best correlation with the measured variables after treatment. DAS4 shows the largest effect size difference from the original DAS1 (0.48).

	DAS1 ES	DAS2 ES	DAS3 ES	DAS4 ES
TJC				
Baseline	0.351**	0.179	0.397**	0.301**
After 6 mo	–0.034	–0.106	-0.067	–0.145
SJC				
Baseline	0.230*	0.114	0.345**	0.266**
After 6 mo	–0.191	–0.309**	–0.157	–0.157
PGA				
Baseline	0.180	–0.072	0.169	–0.043
After 6 mo	–0.484**	–0.666**	–0.110	–0.333**
Pain score				
Baseline	0.176	–0.059	0.162	–0.024
After 6 mo	–0.499**	–0.675**	–0.129	–0.345**
PhGA				
Baseline	0.111	–0.131	0.321**	0.104
After 6 mo	–0.624**	–0.766**	–0.564**	–0.606**
ESR				
Baseline	0.05	–0.300**	0.330**	–0.071
After 6 mo	–0.551**	–0.668**	–0.411**	–0.455**
CRP				
Baseline	0.011	–0.072	0.220*	0.215
After 6 mo	–0.501**	–0.746**	–0.467**	–0.629**

TJC: tender joint count, SJC: swollen joint count, PGA: patient global health assessment, PhGA: physician global health assessment, ESR: erythrocyte sedimentation rate, CRP: C-reactive protein. * $p < 0.05$; ** $p < 0.01$.

joint counts, physician global assessment, and CRP. The scoring system suggested in this study was found to yield the best target value for therapeutic efficacy, in contrast with the traditional DAS and other tested DAS scores. Analyzing the mean values of the DAS2, DAS3, and DAS4 in relation to patients who were found to be improved and not improved using the

Table 5. Effect size and effect size difference of different disease activity scores in comparison to DAS1.

	Effect Size (95% CI)	Effect Size Difference (95% CI)	p
DAS1	1.74 (1.56–1.92)	—	
DAS2	2.09 (1.86–2.32)	–0.35 (–0.48 – (–0.22))	< 0.001
DAS3	1.87 (1.69–2.05)	–0.13 (–0.21 – (–0.04))	< 0.01
DAS4	2.22 (2.0–2.44)	–0.48 (–0.62 – (–0.33))	< 0.001

Table 6. Effect size and effect size difference of ESR and CRP.

	Effect Size (95% CI)	Effect Size Difference (95% CI)	p
ESR	1.08 (0.89–1.27)	—	
CRP	1.03 (0.83–1.23)	0.053 (–0.13–0.24)	NS

NS: nonsignificant.

Table 7. Correlation of age and duration of illness with effect size of different DAS scores studied, CRP, and ESR, and effect size difference between ESR and CRP.

	Age		Duration of Illness	
	r	p	r	p
Effect Size				
DAS1	–0.127	NS	–0.452	< 0.01
DAS2	0.143	NS	–0.363	< 0.01
DAS3	–0.100	NS	–0.418	< 0.01
DAS4	0.148	NS	–0.269	0.057
CRP	–0.036	NS	0.095	NS
ESR	–0.198	NS	–0.346	< 0.05

NS: nonsignificant.

original DAS, the DAS4 presented the highest lower limit (2.76). Few studies have examined changes in these measures in correlation with the DAS scores, and most of the findings were reported as subsidiary notes in the clinical trials. In a study to assess the efficacy of leflunomide compared to MTX and placebo[12], the authors noted that in a subgroup of patients with RA, when TJC and SJC measures were improved, patients' self-reports of pain and functional disability did not improve significantly. In another study to assess the ACR scoring system, Albert, *et al*[13] recommended an alteration in the traditional ACR response, and showed that modification of the ACR response criteria discriminated responders from nonresponders to drug therapy better than the traditional ACR response.

Laboratory assessment of inflammatory markers remains a controversial issue. In some studies, a normal ESR was seen in up to 40% of patients with RA in their first visit. Other studies noted that ESR tends to be stable over the longterm course of RA[13,14]. On the other hand, Ward[15] showed that ESR was more sensitive to change than CRP at 12 and 24 weeks of treatment. Results of our study showed that ESR and CRP differed in their relation with duration of illness. Also, when CRP was used instead of ESR in the assessment of disease activity, DAS2 performed better than DAS1, suggesting that the CRP conveys more accurate information than the ESR. However, when assessed in terms of sensitivity to monitor changes of disease activity, there was no significant difference between ESR and CRP. These findings agree with other studies that compared ESR and CRP and showed that they performed similarly in detecting patients with RA who met the preliminary ACR improvement criteria[16]. Another study showed that, where high baseline CRP concentrations were reported, the sensitivity to change of the CRP was higher[17]. This was supported by the more recent ASPIRE study[18], which showed that high CRP/ESR may identify patients for whom early treatment with biologic therapy was appropriate. Moreover, the ASPIRE study reported that baseline CRP predicted the radiological benefit in patients with RA who were treated with infliximab and MTX. Such results might be explained by the findings that ESR is influenced by the hemoglobin concentration and by serum levels of immunoglobulins and rheumatoid factor[19]. Therefore, because of these associations, ESR may be considered a less specific measure of the acute phase response than CRP. The negative point in assessment of CRP is that there is no standard method for CRP measurement. This may vary between different laboratory methods such as immunodiffusion, nephelometry, fluorescent immunoassay, and rocket immunoelectrophoresis.

Our results showed that while both the patient global assessment and the pain score were significantly correlated with acute phase reactants (ESR and CRP) after 6 months of therapy, neither was correlated with the TJC. In contrast, the physician global assessment correlated significantly with all these factors, including patient global assessment and pain score. Interestingly, the highest correlation was found between patient global assessment and pain score. This raises the point that indices such as patient global assessment and pain score might be ambiguous as measures of disease activity in patients with RA. It correlates more with patient perception and other self-report indices than other objective activity measures. In some studies, it was reported that global scores were unchanged despite significant change in disease activity[20,21]. It has been reported that indicators such as patient global assessment and pain score are affected by many factors other than systemic inflammation[22]. These include age, sex, psychological status, rheumatological examination, and functional disability. Moreover, mechanical or degenerative changes may be the cause of pain and/or functional disability. However, most patients cannot differentiate between mechanical causes and the inflammatory process, hence they consider this as a part of the disease activity process, although such factors can be addressed using assistive devices or with ortho-

pedic surgery. Further, even the outcome of orthopedic surgery may affect scoring of these indicators. Findings similar to our results were reported in a recent study[23] to assess the validity of self-report indices (i.e., the Bath Ankylosing Spondylitis Disease Activity Index) as an indicator of disease activity in patients with psoriatic arthritis. The authors found that self-report indices are inadequate on their own as indicators of disease activity, and that other measures such as joint count and physician assessment were the most important indicators of disease activity.

Our finding of negative correlation between duration of illness and the effect size of treatment agrees with recent reports that delaying treatment results in more disease activity, greater joint damage, and more physical disability[18,24,25]. This can be explained by the results of other studies, which found that patients with early disease have rapid progression, with erosive damage present as early as 4 months[26-28]. Within 2 years, up to 89% of the patients may develop erosions[29]. This provides evidence that there is a critical window of opportunity within the first period of disease onset when the rate of radiographic progression can be reset. These findings may also explain the discrepancy noted between the patient global assessments and pain scores reported by patients and the other disease activity indicators.

In summary, it appears likely that self-report indices such as patient global assessment and pain score are inadequate indicators on their own of disease activity. Joint counts and physician ratings are likely to be the most important measures of disease activity. Caution should be taken when considering these self-report indices, as they may be measuring a somewhat different construct than active inflammation. These data raise the concern that potential control of inflammation according to the currently used measures of disease activity in a short-term clinical trial may not translate into optimal clinically adequate effectiveness. Assessment of disease activity using the Disease Activity Score might profitably be amended by using the physician global assessment and CRP level in the disease activity score equation. Such modification is important for quantifying the degree of improvement of patients undergoing a specific therapeutic modality and for comparing between different therapeutic modalities. However, it is our recommendation that all 7 criteria measured in this study, in addition to the Health Assessment Questionnaire, be routinely documented at each patient's visit, since they fulfil different functions. This approach would generate a more comprehensive picture of the efficacy of any particular therapeutic modality.

REFERENCES

1. Lansbury J. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis: theoretic and clinical considerations. Arthritis Rheum 1958;1:505-22.
2. Ritchie DM, Boyle JA, McInnes JM. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. QJM 1986;37:393-406.
3. Felson DT, Anderson JJ, Boers M. 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.
4. Van der Heijde DMFM, van't Hof M, van Riel PLCM, van de Putte LBA. Validity of single variables and indices to measure disease activity in rheumatoid arthritis. J Rheumatol 1993;20:538-41.
5. Van der Heijde DMFM, van't Hof M, van Riel PLCM, van de Putte LBA. Development of a disease activity score based on judgement in clinical practice by rheumatologists. J Rheumatol 1993;20:579-81.
6. Van der Heijde DMFM, van't Hof M, van riel PLCM, Rijswijk MH, van de Putte LBA. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. Ann Rheum Dis 1992;51:177-81.
7. Pincus T, Sokka T. Partial control of core data set measures and Disease Activity Score (DAS) measures of inflammation does not prevent long term joint damage: evidence from longitudinal observations over 5-20 years. Clin Exp Rheumatol 2002;20 Suppl 27:S42-7.
8. Arnett FC, Edworthy SM, Bloch DA et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315-24.
9. RheDASLite. Available from: http://www.rhedas.net/rhedaslite/rhedaslite.php. Accessed January 26, 2006.
10. Bellamy N. Clinimetric concepts in outcome assessment: the OMERACT filter. J Rheumatol 1999;26:948-50.
11. Ward MM. Clinical measures in rheumatoid arthritis: which are most useful in assessing patients? J Rheumatol 1994;21:17-27.
12. Anderson JJ, Felson DT, Meenan RF, Williams HJ. Which traditional measures should be used in rheumatoid arthritis clinical trials? Arthritis Rheum 1989;32:1093-9.
13. Albert DA, Huang G, Dubrow G, Colleen BM, Berlin JA, Williams HJ. Criteria for improvement in rheumatoid arthritis: Alternative to the American College of Rheumatology 20. J Rheumatol 2003;31:856-66.
14. Wolfe F, Michaud K. The clinical and research significance of the erythrocyte sedimentation rate. J Rheumatol 1994;21:1227-37.
15. Ward MM. Relative sensitivity to change of the erythrocyte sedimentation rate and serum C-reactive protein concentration in rheumatoid arthritis. J Rheumatol 2003;31;884-95.
16. Paulus HE, Ramos B, Wong WK, et al. Equivalence of the acute phase reactants C-reactive protein, plasma viscosity, and Westergren erythrocyte sedimentation rate when used to calculate American College of Rheumatology 20% improvement criteria or the Disease Activity Score in patients with early rheumatoid arthritis. J Rheumatol 1999;26:2324-31.
17. Wolfe F. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. J Rheumatol 1997;24:1477-85.
18. St. Clair EW, van der Heijde DM, Smolen JS, et al. Active Controlled Study of Patients Receiving Infliximab for the Treatment of Rheumatoid Arthritis of Early Onset Study Group. Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. Arthritis Rheum 2004;50:3432-43.
19. Wolfe F. Comparative usefulness of C-reactive protein and erythrocyte sedimentation rate in patients with rheumatoid arthritis. J Rheumatol 1997;24:1477-85.
20. Kaye JJ, Fuchs HA, Moseley JW, Nance EP Jr, Callahan LF, Pincus T. Problems with the Steinbrocker staging system for radiographic assessment of the rheumatoid hand and wrist. Invest Radiol

1990;25:536-44.
21. Hochberg MC, Chang RW, Dwosh I, Lindsey S, Pincus T, Wolfe F. The American College of Rheumatology 1991 revised criteria for the classification of global functional status in rheumatoid arthritis. Arthritis Rheum 1992;35:498-502.
22. Treves R, Scotto C, Bertin P, Bonnet C. Do visual analogue scale pain scores change during a rheumatologist visit. Joint Bone Spine 2002;69:234-5.
23. Taylor WJ, Harrison AA. Could the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) be a valid measure of disease activity in patients with psoriatic arthritis? Arthritis Care Res 2004;51:311-5.
24. Maini R, St. Clair EW, Breedveld F. Infliximab (chimeric anti tumor necrosis factor-α monoclonal antibody) versus placebo in rheumatoid arthritis patients receiving concomitant methotrexate: a randomized phase III trial. Lancet 1999;354:1932-9.
25. Lipsky PE, van der Heijde DMFM, St. Clair EW, Furst DE, Breedveld F, Kalden J. Infliximab and methotrexate in the treatment of rheumatoid arthritis. N Engl J Med 2000;34:1594-602.
26. Tsakonas E, Fitzgerald AA, Fizcharles MA, et al. Consequences of delayed therapy with second-line agents in rheumatoid arthritis: a 3-year followup on the hydroxychloroquine in early rheumatoid arthritis (HERA) study. J Rheumatol 2000;27:623-9.
27. Lard LR, Visser H, Speyer I, et al. Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. Am J Med 2001;111:446-51.
28. Mottonen T, Hannonen P, Korpela M, et al. FIN-RACo Trial Group. FINnish Rheumatoid Arthritis Combination Therapy. Delay to institution of therapy and induction of remission using single-drug or combination-disease-modifying antirheumatic drug therapy in early rheumatoid arthritis. Arthritis Rheum 2002;46:894-8.
29. Fex E, Jonsson K, Johnson U, Eberhardt K. Development of radiographic damage during the first 5-6 years of rheumatoid arthritis. A prospective follow-up study of a Swedish cohort. Br J Rheumatol 1996;35:1106-15.