EXTENDED REPORT

Comparison of different definitions to classify remission and sustained remission: 1 year TEMPO results

D van der Heijde, L Klareskog, M Boers, R Landewé, C Codreanu, H D Bolosiu, R Pedersen, S Fatenejad, for TEMPO Investigators

1ST **ONLINE**

See end of article for authors' affiliations

Correspondence to:
Professor D van der Heijde, Rheumatology Department, University Hospital, Maastricht, PO Box 5800, 6202 AZ Maastricht, P Debyelaan 25, 6229 HX Maastricht, The Netherlands; dhe@sint.azm.nl

Objective: To assess methods to calculate achieving and sustaining remission in a double blind randomised trial in patients with RA who received etanercept, methotrexate, or an etanercept/methotrexate combination.

Methods: Remission was defined as DAS <1.6, DAS28 <2.6, and ACR70 response. Sustaining remission was analysed in three ways: (*a*) analysis of sustained DAS remission, DAS28 remission, or ACR70 response continuously for 6 months; (*b*) analysis of sustained remission appraised through a continuity rewarded scoring system, which is the weighted sum of all intervals in the study in which patients are in DAS or DAS28 remission; or (*c*) longitudinal modelling of remission odds using generalised estimating equations.

Results: Significantly more patients treated with the etanercept/methotrexate combination reached DAS remission (37%) than those treated with either methotrexate (14%) or etanercept (18%) alone (p<0.01). Results for DAS28 and for the ACR70 response were similar. Agreement between DAS remission and DAS28 remission was good, but agreement between either of these and the ACR70 response was less. Patients in DAS or DAS28 remission had a lower level of disease activity (fewer active joints, lower ESR) than those achieving ACR70 response; the converse was seen using pain VAS. The three methods were comparable for sustainability of remission and showed significant advantage for combination therapy, which increased the number and durability of remission periods.

Conclusions: DAS and DAS28 remission results were similar for assessing achieving and sustaining remission in RA, frequently differing from patients classified as ACR70 responders. The three methods of examining duration of remission produced comparable results.

Rheumatoid arthritis (RA) is associated with significant morbidity, mortality, and healthcare expenditure.[1–4] Remission of disease activity has been considered to be the ultimate goal of treatment in RA. A subcommittee of the American Rheumatism Association (ARA) published a set of preliminary criteria for clinical remission as early as 1981, signifying the importance of this end point.[5] These criteria, however, have not been widely used and instead it has been suggested that remission based on the disease activity score (DAS) or on the DAS28 with limited joint counts should be used.[6] Some people suggest the American College of Rheumatology (ACR) 70 as a surrogate for remission, but remission is a definition of a status, whereas ACR70 is an assessment of response, regardless of the final value. An important aspect in the definition of remission is time: being in remission is much more important if it is sustained over a period of time.

Since the early 1980s there have been significant advances in treatment paradigms for RA, including the use of pulse weekly methotrexate (MTX) and anti-tumour necrosis factor agents.[7 8] Despite these advances, a universally accepted method of assessing remission in RA has not been defined. One possible explanation is that traditional agents are either too toxic or induce remission in too few patients to allow for a meaningful comparison of methods; thus, achievement of remission was considered to be too high a hurdle to overcome for the vast majority of patients.

At present, three composite end points, which are reported to perform similarly, are widely used to evaluate disease activity: the ACR response criteria (ACR20, ACR50, and ACR70), the DAS, and the DAS28.[9–11] Owing to difficulties in application of the ACR remission criteria (inclusion of fatigue and a time requirement of 2 months), most recent trials have used a DAS based definition. These have been shown to correlate well with the ACR definition,[6] although recent findings indicate that the DAS28 definition may be too ''lenient'' (that is, closer to low disease activity than remission).[12] Although ACR response criteria have not been used to define remission, achieving an ACR70 response has been used as an indicator of significant control of disease activity.[13]

In addition to defining a specific cut off point required to achieve remission, the durability of remission is important and needs to be examined. The draft criteria proposed by the ARA included a requirement for a minimum of 2 months in remission[5]; however, this duration was somewhat arbitrary and chosen primarily because 90% of patients considered to be in remission fulfilled this criterion.[5 13] The American Food and Drug Administration (FDA) requires maintaining an

Table 1 Joints assessed for DAS, DAS28, and ACR response in the TEMPO trial

Joint	ACR response	DAS	DAS28
Temporomandibular joints (n = 2)	X	X*	
Sternoclavicular joints (n = 2)	X	X	
Acromioclavicular joints (n = 2)	X	X	
Shoulders (n = 2)	X	X	X
Elbows (n = 2)	X	X	X
Wrists (n = 2)	X	X	X
Metacarpophalangeal joints (n = 10)	X	X	X
Cervical spine (n = 1)	X*	X*	
Distal interphalangeal joints (n = 8)	X		
Hips (n = 2)	X*	X*	
Knees (n = 2)	X	X	X
Ankles (n = 2)	X	X	
Talocalcaneal (n = 2)	X	X*	
Tarsi (n = 2)	X	X*	
Metatarsophalangeal joints (n = 10)	X	X	
Fingers (proximal interphalangeal joints) (n = 10)	X	X	X
Toes (n = 10)	X		

*Counted for pain assessment (Ritchie articular index and ACR response) but not for swelling.

Table 2 Concordance between DAS remission, DAS28 remission, and ACR70 response at week 52

ACR70	DAS remission	DAS baseline ≤ median	DAS baseline > median	Total
No	No	121	167	288 (57%)
No	Yes	32	13	45 (9%)
Yes	No	33	43	76 (15%)
Yes	Yes	70	30	100 (20%)

ACR70	DAS28 remission	DAS baseline ≤ median	DAS baseline > median	Total
No	No	118	170	288 (57%)
No	Yes	35	10	45 (9%)
Yes	No	32	37	69 (14%)
Yes	Yes	71	36	107 (21%)

DAS remission	DAS28 remission	DAS baseline ≤ median	DAS baseline > median	Total
No	No	136	197	333 (65%)
No	Yes	18	13	31 (6%)
Yes	No	14	10	24 (5%)
Yes	Yes	88	33	121 (24%)

ACR70 response for at least 6 months to consider a patient to have achieved a major clinical response.[14] No time component was added to the DAS based remission criteria. A recently proposed method rewards durability of remission by counting the number of periods of remission (length of period defined a priori) and adding extra weight to consecutive periods in remission (continuity rewarded, ConRew score).[15]

This report aims at examining various ways of assessing remission in a comparative, randomised controlled trial (RCT) of a relatively large number of patients treated with highly effective treatments for 1 year. This allowed a meaningful evaluation and comparison of these methods.

METHODS

This study was conducted in accordance with the Declaration of Helsinki. We analysed the data from the TEMPO (Trial of Etanercept and Methotrexate with radiographic Patient Outcomes),[8] which was a multicentre, double blind, parallel design study of adult patients with active RA randomised to one of three treatment groups: etanercept only (25 mg twice weekly), MTX only (7.5 mg escalated to 20 mg weekly), or etanercept and MTX in combination.

Clinical assessments, including complete joint counts, erythrocyte sedimentation rate (ESR), C reactive protein, pain visual analogue scale (VAS), general health VAS, Health Assessment Questionnaire, and physician and patient global assessments, were performed at regular intervals.

Dose administration

In addition to the test drugs, all patients received folic acid supplementation of 5 mg twice weekly. Patients were allowed stable doses of oral corticosteroids (≤10 mg/day of prednisone or equivalent) and one non-steroidal anti-inflammatory drug; disease modifying antirheumatic drugs were not allowed.

Definitions of remission

DAS and DAS28 were calculated as described previously using the Ritchie Articular Index and 44 swollen joint count or the 28 joint count for tender and for swollen joint count, respectively.[9 10] Remission was defined as DAS <1.6 or as DAS28 <2.6. Table 1 lists the specific joints used in calculating the ACR response, DAS, and DAS28. Remission based on the recommendations of the ARA[5] was not

calculated because fatigue was not evaluated in this study. Although the initial ACR response criteria suggested assessment of 68 joints, it was recognised that alternatives could be used, if they are proved to have equal discriminative power.[16] In this study all of the 71 joints assessed were used to calculate the ACR response rates.

Assessment of sustained remission
Continuity rewarded scoring

The extent of continuing remission (based on DAS and DAS28) was appraised by the ConRew scoring system, a weighted sum of DAS and DAS28 remission scores for all intervals in the study.[15] A score of 0 was assigned for non-remission at a given time point and a value of 1 was assigned for remission. The ConRew was changed from 1 to 2 if an interval in remission was followed by a second interval in remission. If a patient was in remission at the final interval, that interval was also given a value of 2. There were 10 intervals of 4 weeks or longer with DAS and DAS28 scores after baseline, leading to a possible maximum ConRew score of 20. One visit 2 weeks after baseline was not included to avoid a potential bias against the MTX treatment group. Comparing the unweighted sum of periods with the ConRew score allows one to distinguish overall length of remission time and durability.

Generalised estimating equations (GEE)

Remission status over time was also evaluated using the generalised estimating equations (GEE) approach for longitudinal models. GEE models adjust for within-patient correlations between time points and produce estimates of time averaged odds ratios (ORs) for binomial data. In addition, GEE models do not require complete data and can be fitted when subjects do not have observations at all time points.[17]

Major clinical response and sustained DAS or DAS28 remission for 6 months

Major clinical response was calculated as an ACR70 response continuously for at least 6 months as suggested by the FDA. For comparison, we also calculated the percentage of patients who maintained DAS or DAS28 remission for 6 months.
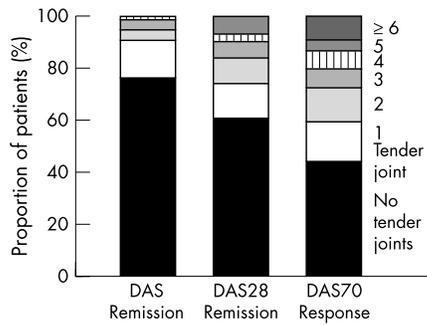
Figure 1 Tender joint count distribution among patients in DAS remission, DAS28 remission, and fulfilling ACR70 response. Mean (SD) tender joint counts for DAS remission, DAS28 remission, and ACR70 responders were 0.4 (1.2), 1.2 (2.4), and 1.8 (2.5), respectively. Medians (25%–75% interquartile range (IQR)) for DAS remission, DAS28 remission, and ACR70 responders were 0 (0–0), 0 (0–2), and 1 (0–3), respectively.

Statistical analysis

Except where otherwise stated, all statistical tests were two sided and significance was assessed at the α = 0.05 level. Unless otherwise mentioned, efficacy analysis used the last observation carried forward (LOCF) imputation approach to include all patients in the time point analysis. ConRew data were appraised by Kruskal-Wallis tests.

RESULTS

A total of 682 patients received the study drugs (MTX = 228, etanercept = 223, etanercept + MTX = 231). Study participants were predominantly women and white with a mean age of 52.9 years. Treatment groups were well balanced for demographic and baseline disease characteristics.[8]

End point criteria

The percentages of patients achieving DAS remission at week 52 were 37%, 18%, and 14% and achieving DAS28 remission were 38%, 18%, and 17%, for the combination, etanercept and MTX, groups, respectively. Combination therapy was significantly better than either of the monotherapies (p<0.01) for both DAS and DAS28 remission. A significantly greater percentage of patients given combination therapy achieved ACR70, compared with the etanercept and methotrexate groups (43%, 24%, and 19%, respectively; p<0.05).

To better understand the relationship between ACR70, DAS remission, and DAS28 remission, and to evaluate the effect of baseline disease activity on this relationship, we
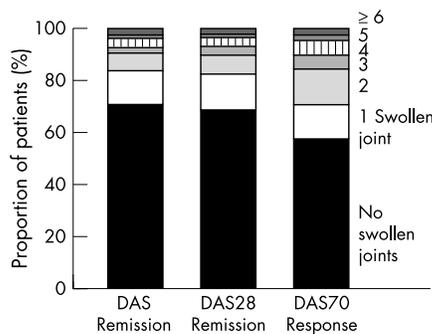

Figure 2 Swollen joint count distribution among patients in DAS remission, DAS28 remission, and fulfilling ACR70 response. Mean (SD) swollen joint count for DAS remission, DAS28 remission, and ACR70 responders were 0.8 (1.6), 0.8 (1.6), and 1.1 (1.8), respectively. Medians (25%–75% IQR) for DAS remission, DAS28 remission, and ACR70 responders were 0 (0–1), 0 (0–1), and 0 (0–2), respectively.
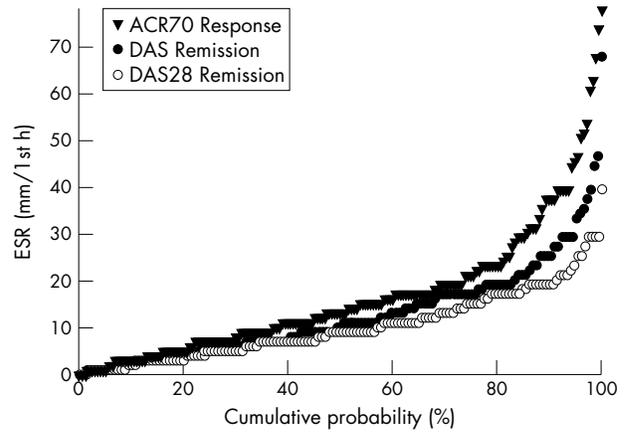

Figure 3 Erythrocyte sedimentation rate (ESR) for individual patients in DAS remission, DAS28 remission, and fulfilling ACR70 response, ordered from the lowest to the highest value and expressed as the cumulative probability. Means (SD) for DAS remission, DAS28 remission, and ACR70 responders were 13.6 (10.5), 11.2 (7.3), and 17.7 (14.4), respectively. Medians (25%–75% IQR) for DAS remission, DAS28 remission, and ACR70 responders were 11 (6–18), 10 (6–16), and 14 (8–22), respectively.
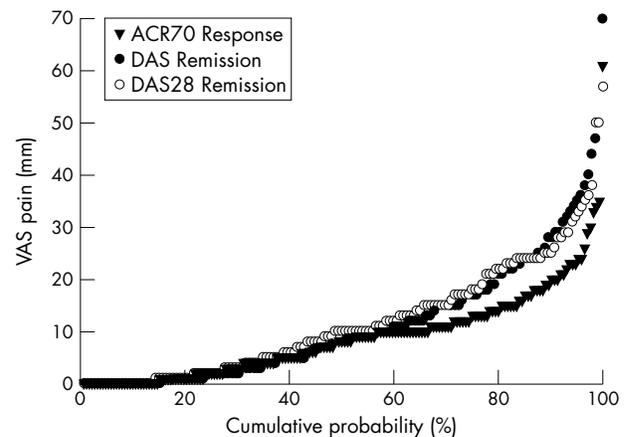

Figure 4 Pain visual analogue scale (VAS) for patients in DAS remission, DAS28 remission and fulfilling ACR70 response, ordered from the lowest to the highest value and expressed as the cumulative probability. Means (SD) for DAS remission, DAS28 remission, and ACR70 responders were 11.5 (12.3), 11.8 (11.3), and 9.0 (8.6), respectively. Medians (25%–75% IQR) for DAS remission, DAS28 remission, and ACR70 responders were 8.0 (2–17), 10 (2–18), and 8 (2–13), respectively.

calculated the concordance rates stratified for baseline level of disease activity (table 2). Overall discordance was greater between either DAS or DAS28 and the ACR70 response, 24% and 22%, respectively, than between DAS and DAS28 (11%). Similarly, concordance, as measured by the κ coefficient, was greater between DAS and DAS28 remission (κ = 0.74) than between the ACR70 response and either DAS or DAS28 remission (κ = 0.45, 0.49, respectively). It was easier to reach an ACR70 response than DAS or DAS28 remission, primarily in patients with high baseline disease activity. In patients with high baseline disease activity discordance between ACR70 and DAS was seen in 56 patients. The majority (43) showed a response according to ACR70 but not according to DAS. By contrast, in the other patients with a lower baseline disease activity, discordance was equally distributed, with roughly 50% meeting either ACR70 or DAS criteria. A similar

Table 3 Unweighted sum scores of 4 week or longer remission periods and ConRew scores for DAS and DAS28

Treatment	No	Unweighted sum score		ConRew score	
		Mean (SD)	Median (range)	Mean (SD)	Median (range)
DAS					
MTX	228	1.0 (2.0)	0 (0–9)	1.7 (3.8)	0 (0–18)
ETN	223	1.1 (2.2)	0 (0–9)	1.9 (4.0)	0 (0–18)
MTX + ETN	231	2.4 (3.1)	1 (0–10)	4.4* (6.0)	1 (0–20)
DAS28					
MTX	228	1.0 (2.0)	0 (0–9)	1.7 (3.6)	0 (0–18)
ETN	223	1.2 (2.3)	0 (0–9)	2.1 (4.2)	0 (0–18)
MTX + ETN	231	2.6 (3.1)	1 (0–10)	4.8* (6.1)	0 (0–20)

*p<0.0001 v MTX and etanercept (ETN) monotherapy.

pattern is seen for the comparison of ACR70 and DAS28. In addition, as expected, we found that patients with low baseline DAS had a better chance of reaching DAS or DAS28 remission than those with high baseline DAS.

We further characterised the disease activity status of patients who were ACR70 responders or were in DAS or DAS28 remission, using four single measures of disease activity at week 52: tender joint counts, swollen joint counts, ESR, and pain VAS (figs 1–4, respectively). From these plots, it is evident that the ACR70 responders included more subjects with a high number of tender and swollen joints and a higher ESR than subjects in DAS or DAS28 remission. The differences between those in remission by DAS and DAS28 were less pronounced, although a higher percentage of patients in DAS remission had no painful joints than those in DAS28 remission. A different pattern was seen when a subjective measure of disease activity, patient's assessment of pain (VAS), was used (fig 4). Patients who were ACR70 responders reported less pain, followed in increasing order by those in DAS remission and DAS28 remission.

Sustained remission criteria
Several methods of assessing remission over a period of time are presented here. The ConRew analysis was used to see if it could produce results that are consistent with other clinical end points in an RCT.[15] Both the unweighted sum scores and the ConRew scores followed a highly skewed distribution, with most patients experiencing only one or a few periods of remission (table 3). Nevertheless, ConRew analysis was able to differentiate between the treatment groups and showed results consistent with those reported for the week 52 time point. Comparing the ConRew scores with the unweighted scores shows that the combination group gained in both number of periods and durability. For example, the MTX group had a mean 1.0 DAS remission periods of 4 weeks or longer, and a mean bonus of 0.7. The results of the etanercept group were only slightly better. In contrast, the combination group had a mean 2.4 periods of remission, and a mean bonus of 2.0 (table 3). Furthermore, the differentiation between the treatment groups using DAS were very close to

that obtained using DAS28, further adding to the validity of this measurement as a way of expressing remission that is persistent over time. Also in this analysis, the scores for ConRew based on DAS28 compared with DAS for each treatment group were higher, indicating that the DAS28 defines remission less strictly.

The major clinical response was also able to differentiate the three treatment groups (table 4). Additionally, there was relatively more separation between the combination therapy and etanercept or MTX monotherapy arms than that reported with the ACR70 at one time point, implying that this measure might be more useful than end point measures for comparing highly effective treatments. When the same 6 month duration of response was applied to DAS and DAS28 remission, similar results were seen for discrimination between the groups, compared with end point criteria (table 4).

A third method examining remission that also includes time was the use of longitudinal modelling, in this case employing GEE. The odds of achieving remission were significantly higher for the combination group than for either of the two monotherapy groups (table 5). To confirm that various methods of imputation for missing information have no notable effect on the outcomes of GEE analysis, we evaluated three different methods for handling missing data and found that in the TEMPO trial equivalent results were obtained regardless of whether (*a*) no data adjustments were made; (*b*) all missing time points were assigned a ''not in remission'' value; or (*c*) an LOCF method was used (table 5).

DISCUSSION
We have closely examined the usefulness of several measures of RA disease activity and disease response in relation to remission using an RCT. We examined two overall approaches to defining remission in an RCT: remission achieved at one time point, cross sectional (in this case, week 52 time point) versus remission criteria that incorporate time as a factor. The cross sectional criteria were further analysed to identify the concordance/discordance rates, the contribution of baseline disease activity to these measures, and the characteristics of patients satisfying each definition

Table 4 Number (%) of patients achieving sustained response over 6 months

Variable	Methotrexate (n = 228)	Etanercept (n = 223)	Etanercept + methotrexate (n = 231)	p Value
Major clinical response*	13 (5.7)	22 (9.9)	55 (23.8)	<0.0001
Sustained DAS remission†	14 (6.1)	13 (5.8)	39 (16.9)	<0.0001
Sustained DAS28 remission‡	8 (3.5)	12 (5.4)	41 (17.9)	<0.0001

*ACR70 response continuously for 6 months; †DAS remission continuously for 6 months; ‡DAS28 remission continuously for 6 months.

Table 5 Longitudinal modelling of remission time-averaged odds ratios (ORs) using generalised estimating equations (GEE) for DAS and DAS28, week 8 through 52

Imputation method	Combination v MTX†	Combination v ETN†	ETN v MTX‡
DAS remission			
None	3.26 (2.15 to 4.93)	2.57 (1.72 to 3.84)	1.26 (0.86 to 1.99)
LOCF	3.37 (2.26 to 5.04)	2.58 (1.75 to 3.79)	1.31 (0.84 to 2.03)
Non-response*	3.49 (2.34 to 5.20)	2.31 (1.58 to 3.37)	1.51 (0.98 to 2.33)
DAS28 remission			
None	3.70 (2.44 to 5.61)	2.56 (1.75 to 3.75)	1.45 (0.92 to 2.27)
LOCF	3.45 (2.36 to 5.06)	2.60 (1.78 to 3.77)	1.33 (0.87 to 2.05)
Non-response*	3.78 (2.62 to 5.46)	2.67 (1.86 to 3.82)	1.42 (0.94 to 2.14)

Results are shown as OR (95% CI).
*Non-response imputation assigns all missing time point values as not in remission; †p<0.0001; ‡not significant.

of remission. Remission that includes a time factor was assessed using three distinct analytical methods to identify reproducibility of outcomes compared with the overall findings of this study.

We found that the two validated cross sectional definitions of disease remission, based on DAS and DAS28, were similar (concordance rate of 88%). However, in most of the comparisons between DAS remission and DAS28 remission, the DAS definition of remission was slightly stricter than the DAS28 criterion. This is also in agreement with a recent publication comparing DAS and DAS28 with the ARA definition of remission.[18] The concordance rate between these measures was not significantly affected by the baseline disease activity level of patients.

We found that the concordance between ACR70 response and DAS (or DAS28) remission was relatively low, especially in patients who have a high level of baseline disease activity. Further characterisation of patients fulfilling one of these three response criteria indicated that patients achieving an ACR70 response at one time point (week 52) were still at a higher level of disease activity at that time point, as assessed by three objective measures of disease activity—swollen and tender joint counts and ESR—than patients achieving a DAS or DAS28 remission. This verifies the notion that the ACR70 response is not a good measure for identifying patients in remission, who by definition should have a very low level of disease activity. The converse was seen when patients' assessment of pain (VAS) was studied. The reason for the discrepancy between the three objective measures and pain VAS is not clear but might be due to a selection of trial patients based on tender and swollen joint count as well as ESR, rather than on pain. A 70% reduction in variables with relatively high values at baseline still results in relatively high values at the end point while meeting the ACR70.

Because RA is a chronic disease that requires treatment over a long period of time, establishment of remission over time is more meaningful than remission at just one time point. We evaluated three methods to assess remission using criteria that incorporate a time factor in defining remission.

ConRew is a method of assessing remission that "awards" points for sustained remission. It does not require patients to be in remission for any prespecified period of time and thus can be applied to studies with short or long duration of treatment. Because the points, both the unweighted sum and the ConRew score, are directly proportional to the—arbitrary—number of visits when disease activity is assessed, the absolute score is meaningless, and could vary from one study to another. This is not a problem within one trial where all groups are assessed with the same frequency. However, before the ConRew score can be usefully applied in clinical care, we need to study what minimum period spent in remission is clinically relevant—that is, translates into better long term outcome—for example, less joint damage or better

function. Rheumatologists will need to agree on the minimum period of remission that can be considered clinically relevant. Until such agreement is reached, the method is useful for comparing treatment arms in one RCT.

In the TEMPO trial, in which a combination of etanercept and MTX resulted in a significantly higher level of disease control, the ConRew data could clearly distinguish this group from the monotherapy arms. Comparing the ConRew score with the unweighted sum score allowed insight into the dynamics of the remission, and showed that the combination group gained in both total remission time and durability of remission. Although the results of this trial for remission experience are notable, our analysis also shows we still have a long way to go: a median of 1 indicates that only 50% of patients in the combination group had one or more remission periods lasting 4 weeks or longer. Disregarding skewness of the mean would indicate that no more than 25% of time in the trial was spent in remission, often in non-contiguous periods.

We also evaluated the usefulness of major clinical response advocated by the FDA. For comparison purposes we applied the same time requirement to DAS and DAS28 remission. All three measures of sustained response could clearly distinguish the combination arm from the two monotherapy arms. As expected, the percentages of patients achieving this level of response were lower than the percentages achieved when response at one time point is the basis for defining remission. It should be noted that use of a prespecified time interval (like 6 months) limits the usefulness of this instrument to studies of longer duration. Additionally, in studies with longer duration of treatment, relatively higher numbers of patients have a chance of achieving a sustained remission compared with those with a shorter duration of treatment. Using a prespecified period of 6 months applied to DAS or DAS28 remission is a less sophisticated method of assessing remission that produces results consistent with ConRew, and can potentially be useful in some clinical studies, despite the mentioned shortcomings.

Finally, we employed a more complex model, GEE, for assessment of remission that includes a time factor in calculating the effect of treatment. This analysis method can be applied to studies of various lengths because it does not require a prespecified duration of remission. It is also a powerful method that does not require imputation for missing data. In the TEMPO trial, the results of this analysis (OR) were similar in each treatment group regardless of the imputation method used, going from one extreme (no adjustment) to another (non-response imputation).

Molenaar et al found that patients in clinical remission, defined according to the modified ARA criteria, still showed radiographic progression during a 2 year follow up, although to a lesser extent than patients having an exacerbation.[19] In fact, they advocated the use of radiographs as one of the

criteria for defining remission. It might be worthwhile to perform a separate remission-type analysis of the 2 year dataset of the TEMPO trial, taking radiographic progression into account.

In conclusion, we provide supportive data from a comparative RCT of RA that compares various ways of defining remission. It was found that in this particular RCT all of these instruments have good discriminative power.

Our data show that DAS remission is the strictest criterion for remission currently used in clinical trials and that an ACR70 response should not be used as a surrogate for remission. We also showed that it is feasible to take time into consideration, leading to even more meaningful data. In this respect, the ConRew method proved to be a valid method, with the advantage that it can be applied to trials of various duration.

ACKNOWLEDGEMENTS

. . . .

Authors' affiliations

D van der Heijde, R Landewé, Rheumatology Department, University Hospital, Maastricht, The Netherlands
L Klareskog, Rheumatology Unit, Department of Medicine, Karolinska Institute/Karolinska Hospital, Stockholm, Sweden
M Boers, Department of Clinical Epidemiology and Biostatistics, VU University Medical Centre, Amsterdam, The Netherlands
C Codreanu, Centrul Metodologic de Reumatologie, Bucuresti, Romania
H D Bolosiu, Clinica Reumatologica, Cluj-Napoca, Romania
R Pedersen, S Fatenejad, Wyeth Research, Collegeville, PA, USA

REFERENCES

1 **Gabriel S**, Tugwell P, O'Brien B, Yelin E, Drummond M, Ruff B, *et al.* Report on the OMERACT task force on economic evaluation. Outcome Measures in Rheumatology. *J Rheumatol* 1999;**26**:203–6.
2 **Cooper NJ**. Economic burden of rheumatoid arthritis: a systematic review. *Rheumatology (Oxford)* 2000;**39**:28–33.
3 **Kvien TK**. Epidemiology and burden of illness of rheumatoid arthritis. *Pharmacoeconomics* 2004;**22**(suppl):1–12.
4 **Navarro-Cano G**, del Rincón I, Pogosian S, Roldán JF, Escalante A. Association of mortality with disease severity in rheumatoid arthritis, independent of comorbidity. *Arthritis Rheum* 2003;**48**:2425–33.
5 **Pinals RS**, Masi AT, Larsen RA. Preliminary criteria for clinical remission in rheumatoid arthritis. *Arthritis Rheum* 1981;**24**:1308–15.
6 **Prevoo ML**, van Gestel AM, van't Hof MA, van Rijiswijk MH, van de Putte LB, van Riel PL. Remission in a prospective study of patients with rheumatoid arthritis. American Rheumatism Association preliminary remission criteria in relation to the disease activity score. *Br J Rheumatol* 1996;**35**:1101–5.
7 **Lee DM**, Weinblatt ME. Rheumatoid arthritis. *Lancet* 2001;**358**:903–11.
8 **Klareskog L**, van der Heijde D, de Jager JP, Gough A, Kalden J, Malaise M, *et al.* Therapeutic effect of the combination of etanercept and methotrexate compared with each treatment alone in patients with rheumatoid arthritis: double-blind randomized controlled trial. *Lancet* 2004;**363**:675–81.
9 **van der Heijde DM**, van't Hof M, van Riel PL, van de Putte LB. Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol* 1993;**20**:579–81.
10 **Prevoo MLL**, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LBA, van Riel PLCM. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a perspective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:44–8.
11 **van Gestel AM**, Anderson JJ, van Riel PLCM, Boers M, Haagsma CJ, Rich B, *et al.* ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. American College of Rheumatology European League of Associations for Rheumatology. *J Rheumatol* 1999;**26**:705–11.
12 **Landewé RB**, van der Heijde D, Voskuyl A, Boers M. Condensed 28-joint counts jeopardize the construct validity of DAS28-remission: a comparison with the original DAS [abstract]. *Arthritis Rheum* 2004;**50**(suppl):S457.
13 **van Riel PLCM**, van Gestel AM. Clinical outcome measures in rheumatoid arthritis. *Ann Rheum Dis* 2000;**59**(suppl 1):i28–31.
14 **FDA**. Rheumatology Working Group of the Medical Policy Coordinating Committee (MPCC) of the Center for Drug Evaluation and Research (CDER), the Center for Biologics Evaluation and Research (CBER), and the Center for Devices and Radiological Health (CDRH). Guidance for industry: clinical development programs for drugs, devices, and biological products for the treatment of rheumatoid arthritis (RA), 1999. http://www.fda.gov/cder/guidance/1208fnl.pdf (accessed 3 August 2005).
15 **Boers M**, Ader H, Bezemer PD, Knol D, Kostense P, Kuik J, *et al.* Modelling 'remission experience' as outcome in a rheumatoid arthritis clinical trial. Exploration of existing and new analytical techniques and graphs [abstract]. *Arthritis Rheum* 2003;**48**(suppl):S612.
16 **Felson DT**, Anderson JJ, Boers M, Bambardier 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.
17 **Liang KY**, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;**73**:13–22.
18 **Fransen J**, Creemers MCW, van Riel PLCM. Remission in rheumatoid arthritis: agreement of disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology (Oxford)* 2004;**43**:1252–5.
19 **Molenaar ET**, Voskuyl AE, Dinant HJ, Bezemer PD, Boers M, Dijkmans BA. Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. *Arthritis Rheum* 2004;**50**:36–42.