

EXTENDED REPORT

Twenty-eight-joint counts invalidate the DAS28 remission definition owing to the omission of the lower extremity joints: a comparison with the original DAS remission

R Landewé, D van der Heijde, S van der Linden, M Boers



Ann Rheum Dis 2006;65:637–641. doi: 10.1136/ard.2005.039859

See end of article for authors' affiliations

Correspondence to:
Dr Robert Landewé,
University Hospital
Maastricht, Department of
Internal Medicine/
Rheumatology, PO Box
5800, 6202 AZ,
Maastricht, Netherlands;
rlan@sint.azm.nl

Accepted 2 October 2005
Published Online First
11 October 2005

Objective: To compare 28 joint disease activity score (DAS28) remission with comprehensive joint count DAS remission in rheumatoid arthritis.

Methods: 620 actually measured paired observations of DAS28 and DAS were analysed in 155 patients. Discordant observations (either DAS or DAS28 below remission cut off level: 1.6 for DAS and 2.6 for DAS28) and concordant observations (both DAS and DAS28 below their remission cut off level) were analysed separately.

Results: 91 of 620 paired DAS observations (15%) were discordant; 87 (in 53 patients) comprised observations in which the DAS28 remission criterion, but not the DAS remission criterion, was met. The reverse was found in only four observations, which were therefore omitted. With the original DAS as standard, DAS28 sensitivity was 95% and specificity 84%. Probability plots showed a swollen joint count >0 in 75% of discordant pairs v 48% of concordant pairs. The same was found for total joint count (TJC) >0 in 90% v 40%; median TJC, 0 v 6) and patient global assessment, but not for ESR. Individual joint analysis showed that 51% of discordant v 18% of concordant observations ($p < 0.0005$) had involvement of lower extremity joints that are not included in the DAS28.

Conclusions: DAS remission is more conservative than DAS28 remission. Activity (tenderness and swelling) in joints not included in the reduced joint counts (ankles, feet) mainly account for the discrepancy between the two assessments. DAS28 remission at a cut off level of 2.6 has insufficient construct validity and should be used with caution in clinical practice and clinical trials.

The original disease activity score (DAS), which was published in 1993 by van der Heijde *et al*, is an instrument to measure the state of disease activity.¹ The DAS includes the following four components: a 44 swollen joint count, the Ritchie articular index,² the erythrocyte sedimentation rate (ESR), and a measure for global health. Cut off points for low and high levels of disease activity have been derived and published, as well as DAS based response criteria.^{3–4} A cut off level of 1.6 was proposed to distinguish patients in remission from those who are not.⁵

For reasons of convenience, a reduced DAS was proposed by Prevoo *et al*, which is based on 28-joint swollen and tender joint counts, and omits grading for tenderness as applied in the Ritchie articular index (DAS28).⁶ Using a conversion formula, the appropriate cut off point for remission measured by DAS28 was reported to be 2.6, and this cut off value was recently validated statistically.⁷

The 28-joint counts differ from the comprehensive joint counts primarily in that they omit the feet and ankle joints. Therefore, there is a possibility that a patient with inflammation only of the feet and ankle joints could classify as being in remission according to the DAS28 remission criterion. We investigated whether such a situation indeed occurs in the context of a clinical trial, by comparing DAS remission and DAS28 remission, as well as their individual components.

METHODS

We investigated patients participating in the COBRA (Combinatietherapie Bij Reumatoïde Artritis) study.^{8,9} COBRA was a 56 week multicentre clinical trial that

randomly assigned 155 patients who met the American College of Rheumatology criteria for rheumatoid arthritis. All patients had early active disease (less than two 2 years, median four months). Previous treatment with antimalarial drugs (but no other disease modifying antirheumatic drugs) was allowed but occurred very infrequently. One group was treated with a combination of sulfasalazine, methotrexate, and initially, high dose oral prednisolone (COBRA regimen), and the other group was treated by sulfasalazine alone. The prednisolone dosage was 60 mg/day in the first week and tapered in weekly steps to the maintenance dosage of 7.5 mg/day in week 7. Prednisolone and methotrexate were tapered and stopped after weeks 28 and 40, respectively, while sulfasalazine was continued.

Patients were assessed frequently with unequal time intervals. For the analysis in this study, we made use of data obtained at baseline and at three month, six month, nine month, and 12 month time points.

Disease activity was measured at all time points by trained research nurses who were not involved in the decision regarding continuation or change of study treatment or other treatment decisions. Disease activity was measured using the disease activity measures of the World Health Organisation/International League of Associations for Rheumatology core set.¹⁰ The measures relevant to this study included a tender joint count (TJC, 68 joints), a swollen joint count (SJC, 44 joints); the Ritchie articular index (2), patient global

Abbreviations: ACR, American College of Rheumatology; COBRA, Combinatietherapie Bij Reumatoïde Artritis; DAS28, 28 joint disease activity score; SJC, swollen joint count; TJC, tender joint count

Table 1 DAS28 remission in case of DAS remission in 620 observations in 156 patients who participated in the COBRA trial

	DAS remission present	DAS remission absent	Total
DAS28 remission present	77	87	164
DAS28 remission absent	4	452	456
Total	81	539	620

Sensitivity = 95%, specificity = 84%, positive predictive value = 47%; negative predictive value = 99%.
DAS, disease activity score; DAS28, 28 joint disease activity score.

assessment of wellbeing (10 cm visual analogue scale with worst (10) and best (0) imaginable health status as extremes); and ESR (Westergren method).

The DAS and the DAS28 were computed using the formulae below. For inferring the DAS28, the 44 swollen joint count and the 68 tender joint count were recalculated into 28 joint counts by using the source data. DAS remission was considered to be present if DAS was <1.6 and DAS28 remission if DAS28 was <2.6.

$$\text{DAS} = 0.54\sqrt{\text{RAI}} + 0.06 \text{SJC} + 0.33 \text{Ln}(\text{ESR}) + 0.007(\text{GH})$$

$$\text{DAS28} = 0.56\sqrt{\text{TJC28}} + 0.28\sqrt{\text{SJC28}} + 0.70 \text{Ln}(\text{ESR}) + 0.014(\text{GH})$$

where RAI = Ritchie articular index, SJC = swollen joint count, GH = global health (10 cm visual analogue scale with 0 = best and 10 = worst), and Ln = natural logarithm.

Analysis

For this study, all observations per patient were counted as independent observations. The focus of the study was on discordance with regard to DAS remission and DAS28 remission. A discordant observation was defined as an observation in which the DAS remission criterion was met while the DAS28 remission criterion was not met, and vice versa. Subsequently, we investigated the swollen joint count, the tender joint count, ESR, and the patient global assessment in discordant remission pairs, and compared them with concordant remission pairs. For this purpose, probability plots were constructed. Probability plots visualise every single observation by plotting them in order from the lowest to the highest value.¹¹ Unlike descriptive statistics, probability plots have the advantage of showing the coherence of the dataset. We also calculated means and centiles of swollen joint count, tender joint count, ESR, and patient global assessment, and we compared swelling in lower extremity joints (ankles, feet) in concordant and discordant pairs, using source data.

RESULTS

We investigated 620 paired observations of DAS and DAS28 in 155 different patients. Table 1 shows the concordance and discordance rates of these observations.

In 168 of 620 observations (27%), the remission criterion was met according to at least one of both remission definitions. In only 77 observations (13%) in 35 patients was the remission criterion met by both definitions, and in 91 observations (15%) in 53 patients paired observations were discordant with respect to remission. In 87 of 91 discordant observations in 49 patients (96%) DAS28 remission but not DAS remission was met. DAS remission but not DAS28 remission was met in only four of 91 observations in four patients (4%). Subsequently, comprehensive swollen and tender joint counts, ESR, and patient global assessments were investigated in the discordant pairs of observations with remission. Because DAS28 remission in the absence of DAS

remission was extremely rare ($n = 4$), we omitted this type of discordance from further analysis.

If the more conservative DAS remission was considered the gold standard, the sensitivity of DAS28 remission was 95%, the specificity 84%, the positive predictive value (PPV) only 47%, and the negative predictive value (NPV) 99%. Attempts to improve the statistical performance of DAS28 by changing the cut off level for remission using receiver operating characteristics (ROC) analysis resulted in an optimal (highest agreement) cut off level for DAS28 of 2.73 (sensitivity 99%, specificity 81%, PPV 43%, NPV 100%). Forcing a cut off level towards increased specificity increases PPV, but at the cost of sensitivity: a cut off level for DAS28 remission of 2.0 resulted in a PPV of 66% and a sensitivity of 70%.

The probability plots in fig 1 represent individual observations with DAS28 remission, stratified by the presence of DAS remission. It is obvious that discordant observations of remission (fulfilling DAS28 remission but not DAS remission) include far higher swollen or tender joint counts as well as patient global assessments that are worse than concordant observations (both DAS28 remission and DAS remission), pointing to a higher level of disease activity, with many active joints. The median (50th centile) number of swollen joints of 44 possibly swollen joints was 0 in the concordant observations, as compared with 3 in the discordant observations. The median number of tender joints was 0 in concordant observations, as compared with 6 in the discordant observations.

The probability plot shows that a swollen joint count >0 was found in 75% of discordant pairs *v* 48% of concordant pairs. Similarly, a tender joint count >0 was found in 90% of discordant *v* 40% of concordant observations.

The discrepancy with regard to ESR was only minimal: concordant and discordant observations of remissions had similar ESR values, and high ESRs (>30 mm) were only seen in the discordant pairs.

The entire plot for the patient global assessment is shifted to the left for the discordant *v* concordant pairs, indicating a higher overall assessment (worse status). In line with this, we found in the discordant observations that patient global assessment correlated better with the 44 joint SJC ($r = 0.33$) than with the 28 joint SJC ($r = 0.03$), and with the 68 joint TJC ($r = 0.48$) than with the 28 joint TJC ($r = 0.23$), suggesting significant disease activity in the joints omitted by the 28 joint count.

Subsequently, we addressed the hypothesis that DAS28 remission could occur in the absence of swelling in the 28 joint count but in the presence of swelling in the joints not belonging to the 28 joint count. Table 2 shows that in observations with DAS28 remission based on an SJC of zero joints, important and clinically relevant swelling could occur in residual joints. The same was found with respect to tenderness. Table 2 also provides the mean SJC and TJC for the different types of joint count (comprehensive, condensed, and residual) for different scenarios. Importantly, in the scenario with discordant remission, the residual SJC accounted for 59% of the total SJC, where it was much lower in the other scenarios (and should be 16 of 44 (36%) assuming a proportional distribution). The residual TJC even accounted for 78% of the total TJC, where it was much lower in the other scenarios (and should be 40 of 68 (59%) assuming a proportional distribution). These results suggest that in a scenario defined by the presence of DAS28 remission but the absence of DAS remission, substantial disease activity (swelling and tenderness) is present preferentially in those joints that are not captured by the condensed 28 joint counts. Note that the mean joint count—but not the ESR—in the discordant remission scenario outweighs the mean joint count in the concordant remission scenario, in agreement with the probability plots in fig 1. Note also that this discordant remission scenario is the only one in which mean

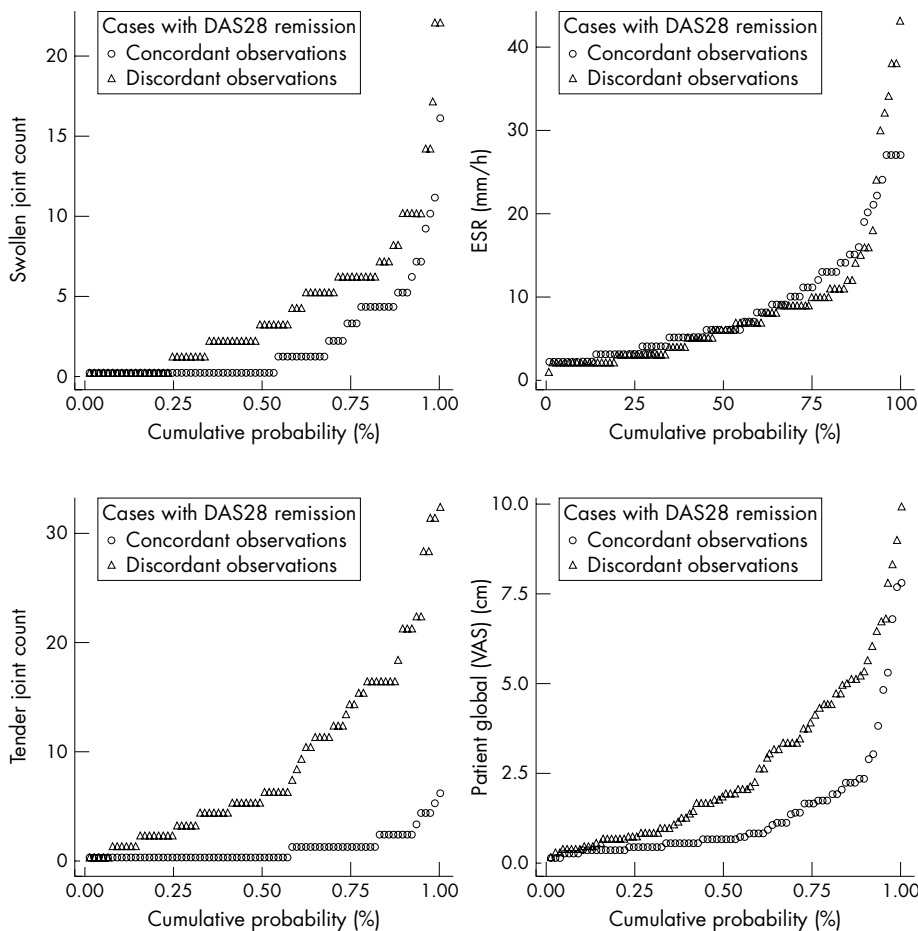


Figure 1 Probability plots of swollen joint count, tender joint count, erythrocyte sedimentation rate (ESR), and patient global assessment of wellbeing in observations with DAS28 remission, with (concordant) or without (discordant) DAS remission. DAS, disease activity score; DAS28, 28 joint disease activity score.

DAS is numerically higher than mean DAS28—another argument for an underestimation of disease activity by DAS28 in this particular scenario.

We subsequently analysed whether the discrepancies between DAS remission and DAS28 remission could be specifically attributed to involvement of ankles and feet.

Seventy one of 77 concordant observations and 83 of 87 discordant observations could be evaluated for swelling of lower extremity joints using the source data. Thirteen of 71 concordant observations (18%) v 42 of 83 discordant observations (51%) ($\chi^2 = 16$; $p < 0.0005$) had swelling in at least one joint of the lower extremities (range 1 to 4).

Table 2 Swelling and tenderness as derived from different joint counts for different scenarios of disease activity, based on DAS28 and DAS remission criteria

Variables	DAS28: No remission DAS: No remission	DAS28: Remission DAS: No remission (Discordant remission)	DAS28: Remission DAS: Remission (Concordant remission)
Swollen joint count from the 44 joint count	10.7 (7.7)	4.1 (4.5)	1.8 (3.0)
Swollen joint count from the 28 joint count	8.0 (6.2)	1.7 (2.3)	1.2 (2.2)
Residual swollen joint count (from 16 joints)*	2.7 (4.2)	2.4 (3.9)	0.6 (1.4)
Residual swollen joint count if SJC28=0	8.6 (10.5)	2.6 (5.2)	0.4 (1.1)
Tender joint count from the 68 joint count	17.9 (14.8)	8.8 (8.0)	0.8 (1.3)
Tender joint count from the 28 joint count	8.7 (7.1)	1.9 (1.9)	0.4 (0.7)
Residual tender joint count (from 40 joints)	9.3 (9.9)	6.9 (7.2)	0.4 (1.1)
Residual swollen joint count if TJC28=0	16.3 (16.1)	6.5 (9.7)	0.4 (0.9)
Erythrocyte sedimentation rate (mm)	30.1 (24.9)	8.1 (8.1)	8.6 (7.0)
Patient global assessment (VAS, mm)	3.5 (2.5)	2.5 (2.3)	1.2 (1.6)
DAS28	4.4 (1.1)	2.1 (0.4)	1.7 (0.5)
DAS	3.6 (1.1)	2.3 (0.7)	1.1 (0.4)

Values are mean (SD).

*The residual swollen (tender) joint count involves joint captured by the comprehensive swollen (tender) joint count, but omitted by the condensed 28 joint swollen (tender) joint count.

DAS, disease activity score; DAS28, 28 joint disease activity score; SJC28, 28 joint swollen joint count; TJC28, 28 joint tender joint count; VAS, visual analogue scale.

DISCUSSION

This study clearly shows that the concepts of DAS remission and DAS28 remission may disguise different situations with respect to disease activity state. It is obvious that DAS remission more closely resembles the situation that clinicians intuitively link with the absence of disease activity than DAS28 remission. Because a swollen joint count of >3 or 4 cannot possibly be compatible with absence of disease activity, it may be concluded that the DAS28 remission (and to a lesser extent DAS remission) is not a true state of remission. We have clearly shown here that concordant observations of remission occur in lower activity states than discordant observations: the joint counts and the patient global—but not the ESR—are much lower in the former than the latter situation. As DAS remission requires low or no tenderness and the absence of swelling in far more joints than DAS28 remission, the explanation for this discrepancy must lie in residual disease activity in joints not captured by the DAS28. Indeed we were able to show that in the absence of swelling or tenderness in the 28 joint count, fulfilling DAS28 remission, there could still be a significant number of observations of swelling or tenderness in one or more residual joints, preferentially in the discordant observations (table 2). Additionally, we were able to show that feet and ankles were far more often involved in discordant observations than in concordant observations. The practicality of the reduced 28 joint count comes at the cost of accuracy. In sum, the DAS28 remission criterion unacceptably overestimates DAS remission, and most probably ACR remission.

Fransen *et al* recently showed in their clinical follow up cohort that, notwithstanding these inherent fallacies of the DAS28, a cut off value of 2.6 appeared to be the best level for remission.⁷ The cut off level of 2.6 was statistically optimal in that it was most accurate in detecting patients with clinical remission according to the ACR remission criteria. The latter comprise a statistical balance between sensitivity and specificity which, however, does not preclude misclassifications in either direction. Both the sensitivity and the specificity of the DAS28 in classifying patients according to the ACR remission criteria were 87% at the optimal cut off level of 2.6. The ACR criteria for remission are stringent in that they require the absence of signs of disease activity, although they allow the presence of activity in one of six items. We agree with Fransen *et al* and many others that the ACR criteria for remission are difficult to handle, but misclassifying 13 of every 100 patients with residual disease activity as being in remission seems unacceptable.

Interestingly, Fransen *et al* also found that the DAS performed better than the DAS28 in detecting remission. They found a greater area under the receiver operating characteristic (ROC) curve for DAS than for DAS28 (both at their optimal cut off points) for detecting ACR remission, confirming DAS remission to be closer to the ACR than the DAS28 remission, and again suggesting lower residual joint counts in DAS remission than in DAS28 remission. We directly visualised residual joint counts by probability plots, and showed that—with a few exceptions—they were very low in observations with DAS remissions but unjustifiably high in a substantial proportion of observations with DAS28 remission but without DAS remission. Our attempts to improve the performance of DAS28 remission showed that a better (higher) cut off level could be obtained by ROC analysis against DAS remission. However, the problem of misclassification based on residual activity in the lower extremity joints only increases, which in turn reflects the limitation of ROC analysis: it provides statistically optimal cut off levels without weighing the direction of misclassification. However, forcing to a lower cut off level increased specificity, but at too high a cost of sensitivity.

An important factor in this regard pertains to our recent observation in the TEMPO trial in which we did not find an important discrepancy between DAS remission and DAS28 remission. This seemingly discrepant observation can be explained by the fact that residual disease activity among patients in remission was much lower in the TEMPO trial than in the COBRA trial.¹²

In the context of the Outcome Measures in Rheumatology (OMERACT) initiative, work has been done to define an upper limit of “minimal disease activity”, a state deemed satisfactory by both patient and physician given current treatment options and limitations. A preliminary set of definitions was accepted at the OMERACT 7 conference.¹³ These were derived from physician opinions of profiles (descriptions of trial patients) and real patients seen in practice. For DAS28, values between 2.60 and 2.85 optimally discriminated between patients considered to have minimal disease activity and those with higher levels of disease activity. Although in the end the value of 2.85 was selected, these findings strengthen our conclusion that a DAS28 level of 2.60 is not sufficiently specific to serve as a cut off point for remission.

A weakness of our analysis could be that we analysed the data as though they were from 620 independent observations. A patient with residual disease activity in the feet and ankles at a particular time point who meets the DAS28 remission criterion will fail the DAS remission criterion. This patient will have an increased chance of showing the same picture one time point later, and we may have overestimated the prevalence of discordant observations by doing this. However, the same analysis restricted to single time points (data not shown) did not essentially change the results, which adds to the validity of the observation.

Another potential limitation is that the prevalence of discordant observations depends on the character of the cohort. If one investigates a trial of rheumatoid patients with a very effective treatment, or low disease activity per se, and as a consequence of that many observations with “deep” remission, the discrepancy between DAS remission and DAS28 remission could disappear, simply because residual disease activity in joints not captured by reduced joint counts is not present. However, we suggest the analyses from our dataset (patients on COBRA treatment) remain relevant today—that is, for the levels of disease activity obtainable with current treatment strategies including biological agents: in a recent head to head trial, COBRA treatment was equivalent to high dose methotrexate plus infliximab in an active early rheumatoid arthritis population.¹⁴

In conclusion, we found the DAS28 definition of remission to be inferior to the original DAS definition in the setting of a head to head treatment trial of early active rheumatoid arthritis (COBRA). This finding may become of particular importance as indices of disease activity are increasingly promoted as tools to “benchmark” the intensity of antirheumatic treatment.¹⁵ As the scenario of an absence of activity in the 28 joint count but the presence of activity in the residual joints in fact exists, we propose that no patient can be classified as being in remission without a full joint assessment.

Authors' affiliations

R Landewé, D van der Heijde, S van der Linden, Department of Internal Medicine/Rheumatology, University Hospital Maastricht, Maastricht, Netherlands

M Boers, Department of Clinical Epidemiology and Biostatistics, VU University Medical Centre, Amsterdam, Netherlands

REFERENCES

- 1 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.

- 2 **Ritchie DM**, Boyle JA, McInnes JM, Jasani MK, Dalakas TG, Grieveson P, *et al*. Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q J Med* 1968;**37**:393–406.
- 3 **van Gestel AM**, Prevoo ML, van't Hof MA, van Rijswijk MH, van de Putte LB, van Riel PL. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. *Arthritis Rheum* 1996;**39**:34–40.
- 4 **van Gestel AM**, Anderson JJ, van Riel PL, 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.
- 5 **Prevoo ML**, van Gestel AM, van THMA, van Rijswijk 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 [see comments]. *Br J Rheumatol* 1996;**35**:1101–5.
- 6 **Prevoo ML**, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:44–8.
- 7 **Fransen J**, Creemers MC, Van Riel PL. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology (Oxford)* 2004;**43**:1252–5.
- 8 **Boers M**, Verhoeven AC, Markuse HM, van de Laar MA, Westhovens R, van Denderen JC, *et al*. Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. *Lancet* 1997;**350**:309–18.
- 9 **Landewé RB**, Boers M, Verhoeven AC, Westhovens R, van de Laar MA, Markuse HM, *et al*. COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. *Arthritis Rheum* 2002;**46**:347–56.
- 10 **Boers M**, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, *et al*. World Health Organization and International League of Associations for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol Suppl* 1994;**41**:86–9.
- 11 **Landewé R**, van der Heijde D. Radiographic progression visualised by probability plots. Presenting data with optimal use of the individual values. *Arthritis Rheum*. 2003, 2004;**50**:699–706.
- 12 **van der Heijde D**, Klareskog L, Boers M, Landewé R, Codreanu C, Bolosiu H, for the TEMPO investigators, *et al*. Comparison of different definitions to classify remission and sustained remission: 1 year TEMPO results *Ann Rheum Dis* 2005;**64**:1582–7.
- 13 **Wells G**, Boers M, Shea B, *et al*. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;**32**:2016–24.
- 14 **Goekoop-Ruiterman YP**, de Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, Hazes JM, *et al*. Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): a randomized controlled trial. *Arthritis Rheum* 2005;**52**:3381–90.
- 15 **Grigor C**, Capell H, Stirling A, McMahon AD, Lock P, Vallance R, *et al*. Effect of a treatment strategy of tight control for rheumatoid arthritis (the TICORA study): a single-blind randomised controlled trial. *Lancet* 2004;**364**:263–9.