

# Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate

G Wells,<sup>1</sup> J-C Becker,<sup>2</sup> J Teng,<sup>2</sup> M Dougados,<sup>3</sup> M Schiff,<sup>4</sup> J Smolen,<sup>5</sup> D Aletaha,<sup>6</sup> P L C M van Riel<sup>7</sup>

<sup>1</sup> Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada; <sup>2</sup> Bristol-Myers Squibb, Princeton, New Jersey, USA; <sup>3</sup> Paris-Descartes University, Medicine Faculty and UPRES-EA 4058, AP-HP, Cochin Hospital, Paris, France; <sup>4</sup> University of Colorado, Denver, Colorado, USA; <sup>5</sup> Second Department of Medicine, Hietzing Hospital, Vienna, Austria; <sup>6</sup> Department of Rheumatology, Medical University of Vienna, Vienna, Austria; <sup>7</sup> University Medical Centre, Nijmegen, The Netherlands

Correspondence to: Professor G Wells, Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, K1H 8M5, Canada; gawells@ottawaheart.ca

Accepted 9 April 2008  
Published Online First  
19 May 2008

## ABSTRACT

**Objective:** To validate and compare the definition of the Disease Activity Score 28 based on C-reactive protein (DAS28 (CRP)) to the definition based on erythrocyte sedimentation rate (ESR).

**Methods:** Data were analysed from two randomised, double-blind, placebo-controlled trials of abatacept of 6-month and 12-month duration in patients with rheumatoid arthritis. European League Against Rheumatism (EULAR) response criteria and the proportion of patients in remission (DAS28 <2.6) based on the two DAS28 definitions were examined. Trends in radiographic progression (erosion score, joint space narrowing score and total score) and physical function (Health Assessment Questionnaire Disability Index (HAQ-DI)) across the EULAR responder states (none, moderate and good) were analysed.

**Results:** There was general agreement in determining the EULAR responder state using both DAS28 definitions ( $\kappa = 0.80$ , 95% CI 0.76 to 0.83). Overall, there was 82.4% agreement on the EULAR response criteria; when disagreements occurred, the DAS28 (CRP) yielded a better EULAR response more often than DAS28 (ESR) (12.6% vs 4.9%, respectively). There was also agreement in determining remission:  $\kappa = 0.69$  (95% CI 0.60 to 0.78). Radiographic progression decreased in patients treated with abatacept across EULAR states (from none to moderate to good) based on both definitions. For patients treated with placebo, the trend was not as pronounced, with radiographic scores higher for moderate vs non-responders. For physical function, similar trends were observed across the EULAR states for both DAS28 definitions.

**Conclusions:** The DAS28 (CRP) has been validated against radiographic progression and physical function. While the DAS28 (CRP) yielded a better EULAR response more often than the DAS28 (ESR), the validation profile was similar to the DAS28 (ESR), indicating that both measures are useful for assessing disease activity in patients with rheumatoid arthritis.

(DAS)<sup>2</sup> are widely used. The DAS index combines information relating to the number of swollen and tender joints, in addition to a measure of general health, and the acute phase response. The DAS28<sup>3</sup> is based on a count of 28 swollen and tender joints, with a score ranging from 0 to 9.4,<sup>4</sup> and can be used to objectively evaluate a patient's response to treatment. An absolute level of disease activity can be selected as a clinically meaningful goal for therapeutic intervention; with a value of  $\leq 3.2$  defined as the threshold for a low disease activity state and  $< 2.6$  as the threshold for remission.<sup>5</sup> Alternatively, the European League Against Rheumatism (EULAR) response criteria combine the DAS28 score at the time of evaluation with the change in DAS28 score between two time points, and enable the user to define improvement or response to treatment.<sup>5</sup> The thresholds for low disease activity and remission and the EULAR response criteria provide a standardised guide on how to interpret the DAS28 scores.<sup>4 6 7</sup> The DAS28 based on erythrocyte sedimentation rate (DAS28 (ESR)) has been extensively validated for its use in clinical trials in combination with the EULAR response criteria.<sup>8 9 10 11</sup>

More recently, an alternative formulation of the DAS28 based on C-reactive protein (DAS28 (CRP)) has been proposed and developed.<sup>12</sup> At the present time, the DAS28 (CRP) is not as well established as the DAS28 (ESR), and its validity is currently only inferred by comparison with the DAS28 (ESR). A comparison of the two DAS28 definitions and a formal validation of the DAS28 (CRP) is necessary so that the clinician or patient will have the same confidence using and interpreting the DAS28 (CRP) that they have come to expect when using the DAS28 (ESR). To properly evaluate the validity of the DAS28 (CRP), a well controlled study with assessment of radiographic progression and physical function is needed in a setting in which a change in patient clinical status has occurred. The ATTAIN (Abatacept Trial in Treatment of Anti-tumor necrosis factor (TNF) Inadequate responders) and AIM (Abatacept in Inadequate responders to Methotrexate (MTX)) trials of the selective costimulation modulator abatacept (which modulates the CD80/CD86:CD28 signal required for



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://ard.bmj.com/info/unlocked.dtl>

Agreement on response criteria in rheumatoid arthritis (RA) has allowed better standardisation and interpretation of clinical trial reports. In particular, American College of Rheumatology (ACR) criteria<sup>1</sup> and the Disease Activity Score

full T cell activation) provide such an opportunity.<sup>13 14</sup> Both studies were well controlled, and the measures required for validation of the DAS28 (CRP) were recorded. Significant improvements in clinical measures of disease activity, physical function and health-related quality of life were observed with active treatment vs placebo in both trials. In addition, radiographic assessments were taken during the AIM trial, and patients treated with abatacept demonstrated significant reductions in structural progression compared with patients treated with placebo.<sup>14</sup> Using data from these two trials, the objective of this investigation was to: (1) compare the DAS28 (CRP) with the DAS28 (ESR) by crossclassifying EULAR response states and the proportion of patients achieving DAS28-defined remission; and (2) to validate the DAS28 (CRP) against assessments of radiographic progression and physical function.

## METHODS

### Datasets

Data from two phase III, multicentre, randomised, double-blind, placebo-controlled trials of abatacept in patients with active RA were used in this analysis. In the 6-month ATTAIN trial, patients with an inadequate response to anti-TNF therapy received abatacept or placebo, plus  $\geq 1$  disease-modifying antirheumatic drug.<sup>13</sup> In the 12-month AIM trial, patients with an inadequate response to MTX received abatacept or placebo, plus MTX.<sup>14</sup>

Data are presented for all patients in the ATTAIN and AIM trials for whom ESR and CRP measurements were available.

### Measures of disease activity

Efficacy assessments were taken on study visit days, prior to infusion. The DAS28 considers 28 tender and swollen joint counts, general health (GH; patient assessment of disease activity using a 100 mm visual analogue scale (VAS) with 0 = best, 100 = worst), plus levels of an acute phase reactant (either ESR (mm/h) or CRP (mg/litre)). DAS28 values were calculated as follows:  $\text{DAS28 (CRP)} = 0.56 \cdot \sqrt{(\text{TJC28}) + 0.28 \cdot \sqrt{(\text{SJC28})} + 0.014 \cdot \text{GH} + 0.36 \cdot \ln(\text{CRP} + 1) + 0.96$ ;  $\text{DAS28 (ESR)} = 0.56 \cdot \sqrt{(\text{TJC28})} + 0.28 \cdot \sqrt{(\text{SJC28})} + 0.014 \cdot \text{GH} + 0.70 \cdot \ln(\text{ESR})$ , where TJC = tender joint count and SJC = swollen joint count.

EULAR response states were classified as follows: good responders were patients with an improvement of  $>1.2$  and a present score of  $\leq 3.2$ ; moderate responders were patients with an improvement of  $>0.6$  to  $\leq 1.2$  and a present score of  $\leq 5.1$ , or an improvement of  $>1.2$  and a present score of  $>3.2$ ; non-responders were any patients with an improvement of  $\leq 0.6$ , or patients with an improvement of  $>0.6$  to  $\leq 1.2$  and a present score of  $>5.1$ .<sup>4</sup> DAS28-defined remission was classified as a score of  $<2.6$ .

Physical function was measured at baseline and every 3 months using the Health Assessment Questionnaire Disability Index (HAQ-DI).<sup>15</sup> In the AIM trial, joints in the hands, wrists and feet were assessed radiographically at baseline and 1 year for changes in erosion score (ES), joint space narrowing (JSN) score and total score (TS) using the Genant-modified Sharp scoring system.<sup>16-19</sup>

### Validation analyses for the disease activity score 28 (CRP)

#### Criterion validity

To assess the extent of agreement between the two DAS28 definitions, the EULAR response criteria based on each definition were calculated by treatment group and for pooled

treatment groups, with crossclassification for each of the datasets. The proportion of patients who achieved remission according to the two DAS28 definitions was calculated for the pooled treatment groups.  $\kappa$  coefficients with quadratic weights were calculated and Bland–Altman plots<sup>20</sup> were constructed.

#### Construct validity

To examine the extent to which the DAS28 (CRP) reflects structural damage and physical function, trends in radiographic progression and HAQ-DI scores across the EULAR responder states based on the DAS28 (CRP) and the DAS28 (ESR) were assessed, and trends based on the two DAS28 definitions were compared.

#### Sensitivity to change

To examine the ability of the two DAS28 definitions to detect a treatment effect, the following measures were assessed:<sup>21</sup> treatment difference, relative percentage improvement, standardised response mean (SRM) and relative efficiency in relation to the tender joint count.

## RESULTS

Of the 258 and 133 patients randomised and treated with abatacept or placebo, respectively, in the ATTAIN trial, 171 patients treated with abatacept and 75 patients treated with placebo had ESR and CRP measurements available for analysis. Of the 433 and 219 patients randomised and treated with abatacept or placebo, respectively, in the AIM trial, 351 patients treated with abatacept and 155 patients treated with placebo had ESR and CRP measurements available for analysis. Baseline demographics and clinical characteristics were generally similar for the abatacept and placebo groups in each study (table 1). Baseline demographics and clinical characteristics were comparable between this subset of patients, and the subset of patients who did not have DAS28 (CRP) and DAS28 (ESR) available (data not shown).

Table 2 shows the crossclassification of the EULAR response criteria based on DAS28 (ESR) and DAS28 (CRP) at 6 months for patients in both treatment groups for the ATTAIN and AIM trials combined. There was general agreement in classifying patients as none, moderate and good EULAR responders using the two DAS28 definitions, with a  $\kappa$  (95% CI) of 0.80 (0.76, 0.83) indicating good agreement. The main diagonal numbers in table 2 indicate where the EULAR response criteria based on DAS28 (ESR) and DAS28 (CRP) are in agreement. Overall there was an 82.4% agreement between the EULAR response criteria based on the two definitions (table 2). When disagreements occurred, the DAS28 (CRP) yielded a better EULAR response state more often than the DAS28 (ESR) (12.7% vs 4.9%, respectively). In one instance, a patient was classified as being a non-responder with DAS28 (CRP) but a good responder with DAS28 (ESR). This patient had a CRP value of 22.4 mg/dl and an ESR value of 4.0 mm/h, corresponding to a DAS28 (CRP) score of 4.3 and a DAS28 (ESR) score of 2.4. The patient subsequently discontinued from the ATTAIN trial due to cholangiocarcinoma, which was the probable cause of the elevated CRP levels.

Similar patterns were observed when the ATTAIN and AIM trials were considered separately, and when treatment groups were considered combined or separately (data not shown). A total of 50/752 (6.6%) patients were classified as having achieved remission according to both DAS28 definitions. A total of 83 (11.0%) patients were classified as having achieved

**Table 1** Baseline demographics and clinical characteristics for all randomised and treated patients with C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR) measurements available

	ATTAIN		AIM	
	Abatacept + DMARDs n = 171	Placebo + DMARDs n = 75	Abatacept + MTX n = 351	Placebo + MTX n = 155
Age, years	52.5 (12.1)	52.5 (12.0)	50.8 (12.7)	49.5 (11.6)
Gender, % female	77.2	80.0	77.2	83.2
Race, % Caucasian	96.5	93.3	87.7	86.5
Disease duration, years	12.0 (8.1)	11.7 (9.0)	8.5 (7.1)	8.9 (7.0)
Tender joints	30.9 (13.4)	31.0 (13.6)	31.1 (12.8)	33.7 (13.9)
Swollen joints	22.6 (10.4)	22.1 (10.1)	21.3 (8.3)	22.3 (8.3)
Erosion score*	N/A	N/A	22.4 (18.3)	22.1 (19.7)
Joint space narrowing score*	N/A	N/A	23.3 (20.1)	23.7 (21.3)
Total score*	N/A	N/A	45.7 (37.4)	45.8 (39.8)
Pain assessment, 100 mm VAS	69.5 (20.5)	68.0 (19.0)	63.1 (21.1)	67.3 (18.9)
Physical function, HAQ-DI	1.8 (0.5)	1.7 (0.6)	1.7 (0.7)	1.7 (0.6)
Patient global assessment, 100 mm VAS	68.2 (19.6)	66.7 (22.0)	63.2 (20.9)	63.8 (20.6)
Physician global assessment, 100 mm VAS	67.7 (17.7)	63.7 (18.2)	69.3 (15.3)	68.2 (15.9)
Levels of CRP, mg/dl	4.4 (3.6)	3.5 (3.3)	3.3 (3.2)	2.5 (2.1)
ESR, mm/h	49.9 (27.7)	43.7 (26.9)	44.1 (23.5)	43.0 (24.5)
DAS28 (CRP)	6.5 (0.9)	6.4 (0.8)	6.4 (0.8)	6.4 (0.8)
DAS28 (ESR)	6.9 (1.0)	6.8 (1.0)	6.8 (0.9)	6.9 (0.8)

Data are mean (SD) unless otherwise stated.

\*Values represent baseline readings for all patients with radiographic data at day 365 (n = 328 for abatacept; n = 137 for placebo). AIM, Abatacept in Inadequate responders to Methotrexate; ATTAIN, Abatacept Trial in Treatment of Anti-TNF INadequate responders; DAS28, 28-joint Disease Activity Score; DMARD, disease-modifying antirheumatic drug; HAQ-DI, Health Assessment Questionnaire Disability Index; MTX, methotrexate; N/A, not applicable; VAS, visual analogue scale.

remission with the DAS28 (CRP) and 56 (7.4%) classified as having achieved remission with the DAS28 (ESR). The  $\kappa$  coefficient (95% CI) was 0.69 (0.60 to 0.78), indicating good agreement. A Bland–Altman plot (fig 1) shows a high degree of agreement between the two DAS28 definitions with most observations of the difference between DAS28 (ESR) and DAS28 (CRP) lying between the mean  $\pm$  2SD (represented in the figure by central and outer horizontal dashed lines, respectively). A total of 3.3% of values fell outside 2SD above the mean difference and 1.5% fell outside 2SD below the mean difference, indicating that when disagreement of two definitions occurred, there were more instances in which DAS28 (CRP) was lower than DAS28 (ESR) compared with instances in which it was higher. The mean difference between DAS28 (CRP) and DAS28 (ESR) is  $-0.38$ , indicating a relatively small tendency for the DAS28 (ESR) to classify patients as having a higher score.

Radiographic progression was assessed in the AIM trial.<sup>14</sup> For both definitions, there were several instances in which

radiographic progression did not improve in a linear fashion across the EULAR responder states. This was particularly apparent in the placebo group, where ES, JSN score and TS improved in a quadratic manner with a sharp improvement for good vs moderate responders but a more modest improvement for moderate vs non-responders, and in the case of the DAS28 (CRP) definition a slight deterioration was observed for moderate vs non-responders (table 3). Similar patterns were observed for ES in the abatacept group and in both groups combined. The JSN score in the abatacept group using the DAS28 (ESR) definition followed a more linear pattern across the EULAR responder states, which in turn led to a linear pattern for the JSN scores in both groups combined and for TS in the abatacept and combined groups. For the DAS28 (CRP) definition, the JSN scores and TS followed the quadratic pattern.

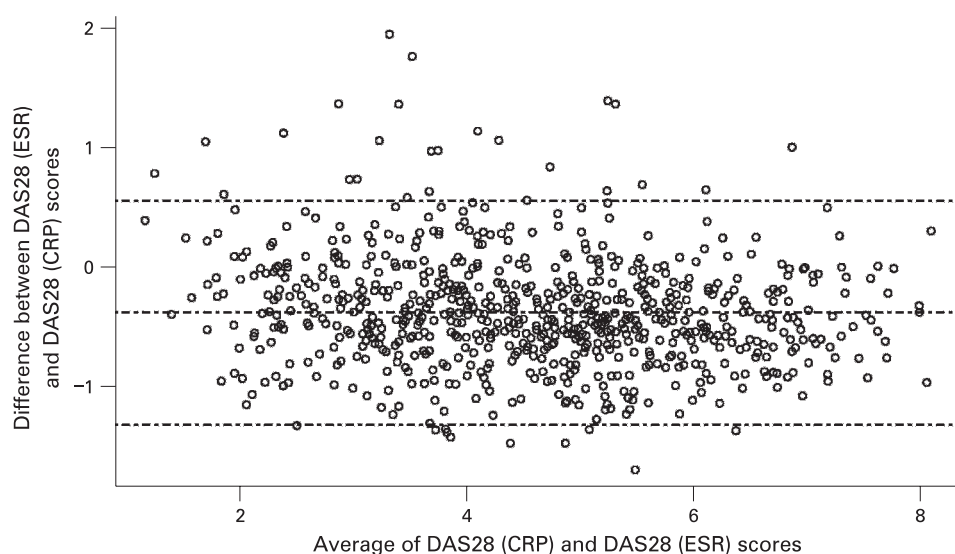
Table 4 and fig 2 show data from the ATTAIN and AIM trials, illustrating the mean improvements from baseline in HAQ-DI at 6 months in patients who were EULAR good,

**Table 2** Crossclassification of patients by the European League Against Rheumatism (EULAR) response criteria based on the 28-joint Disease Activity Score (DAS28) C-reactive protein (CRP) and DAS28 erythrocyte sedimentation rate (ESR) scales at 6 months (AIM (Abatacept in Inadequate responders to Methotrexate) and ATTAIN (Abatacept Trial in Treatment of Anti-TNF INadequate responders) trials, both treatment groups combined\*)

		DAS28 (CRP)		
		Good responder n (%)	Moderate responder n (%)	Non-responder n (%)
DAS28 (ESR)	Good responder n (%)	108 (14.4)†	12 (1.6)	1 (0.1)
	Moderate responder n (%)	56 (7.4)	352 (46.8)†	24 (3.2)
	Non-responder n (%)	0 (0)	39 (5.2)	160 (21.3)†

\*n = 752, based on all randomised and treated patients with CRP and ESR measurements available; †indicates agreement between EULAR response criteria based on DAS28 (ESR) and DAS28 (CRP).

**Figure 1** Bland–Altman plot of 28-joint Disease Activity Score (DAS28) C-reactive protein (CRP) and DAS28 erythrocyte sedimentation rate (ESR) values for patients in the ATTAIn (Abatacept Trial in Treatment of Anti-TNF Inadequate responders) and AIM (Abatacept in Inadequate responders to Methotrexate) trials (both treatment groups) at 6 months. Difference between DAS28 (ESR) and DAS28 (CRP) scores vs mean value of DAS28 (ESR) and DAS28 (CRP) score combined for patients in the ATTAIn and AIM trials (both treatment groups) at 6 months. The central line represents the mean difference between the two measures, and the upper and lower bounds represent 2SD from the mean.



moderate or non-responders, based on the DAS28 (CRP) or DAS28 (ESR) definition. For both definitions, function improved in a linear fashion across the EULAR responder states, and there was agreement on the changes in HAQ-DI scores across the states. For the good and moderate states, there was improvement from baseline in HAQ-DI, and this improvement was more marked in good responders compared with moderate responders. For EULAR non-responders, HAQ-DI scores showed little improvement at 6 months. Similar results were observed for patients in the AIM trial at 1 year. For the combined (abatacept and placebo) group, the mean change from baseline in HAQ-DI was  $-0.92$  and  $-1.03$  for good responders  $-0.62$  and  $-0.64$  for moderate responders and  $-0.27$  and  $-0.27$  for non-responders, for DAS28 definitions based on CRP and ESR, respectively.

Measures to assess the sensitivity of the two DAS28 definitions demonstrate that both have a comparable ability to detect a treatment effect. For the DAS28 (CRP) and DAS28 (ESR), respectively, the treatment difference was  $-18.83$  vs  $-13.22$ ; the percentage improvement was  $-14.42$  vs  $-7.44$ ; the SRM was  $-0.31$  vs  $-0.31$  and the relative efficiency was  $1.93$  vs  $1.97$ .

## DISCUSSION

Using CRP for calculation of the DAS28 is an attractive alternative to ESR for a number of reasons. Firstly, CRP

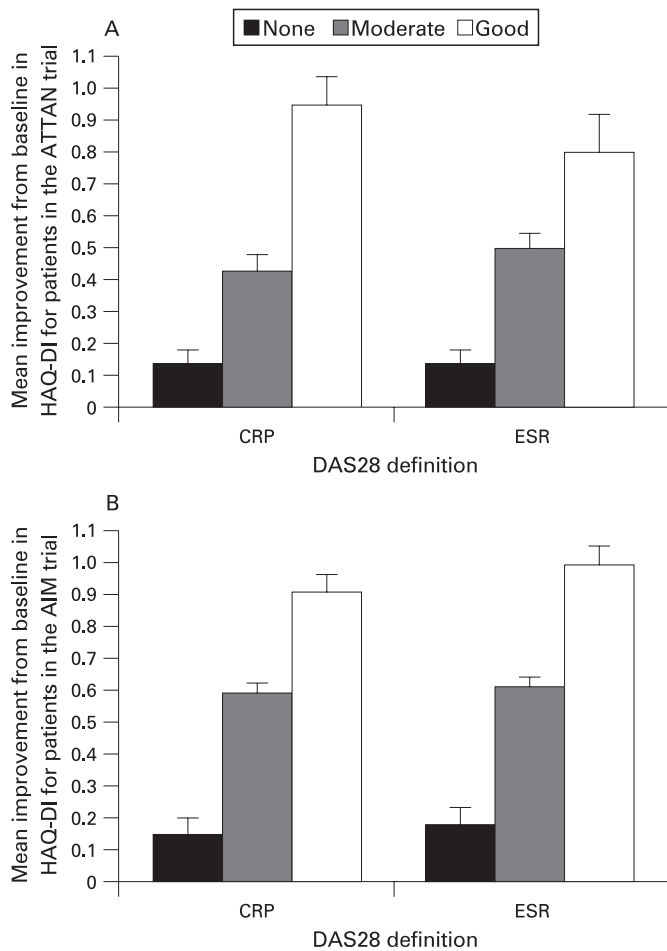
measurements are routinely used in clinical practice, and are often available in circumstances when ESR measurements are not. CRP levels are more sensitive to short-term changes in disease activity,<sup>22</sup> whereas ESR can be influenced by a number of unrelated factors, such as age, gender or plasma proteins. Laboratory tests used to calculate CRP are faster than those used to measure ESR, and measurements can be standardised in a central laboratory for multicentre clinical trials.<sup>7</sup> There is currently less clinical experience using DAS28 (CRP), and a formal validation study with respect to radiographic progression and functional assessment is needed.

Using data from two trials of abatacept, we have compared the DAS28 (CRP) with the DAS28 (ESR), and validated the DAS28 (CRP) against assessments of radiographic progression and physical function. Most patients were classified as having the same EULAR state regardless of which DAS28 definition is used. Despite this, several discrepancies were observed. Firstly, some patients (154/752; 20.5%) were classified as having a lower present DAS28 category when using the DAS28 (CRP) vs the DAS28 (ESR). In 76 out of these 154 cases, this resulted in patients being classified as being in a better EULAR state when measured with DAS28 (CRP). Secondly, certain patients were classified as having either large or moderate improvement, depending on which DAS28 definition was used; this type of discrepancy occurred at a similar frequency with both

**Table 3** Radiographic progression across the European League Against Rheumatism (EULAR) states based on the 28-joint Disease Activity Score (DAS28) C-reactive protein (CRP) and DAS28 erythrocyte sedimentation rate (ESR) values: mean change from baseline to 12 months (AIM (Abatacept in Inadequate responders to Methotrexate) trial)

Radiographic assessment	EULAR response	Abatacept group (n = 328)		Placebo group (n = 137)		Combined group (n = 465)	
		CRP	ESR	CRP	ESR	CRP	ESR
Erosion score	Good	0.37 (1.13)	0.37 (1.17)	0.41 (1.15)	0.10 (0.23)	0.37 (1.13)	0.35 (1.14)
	Moderate	0.72 (1.70)	0.63 (1.56)	1.48 (3.70)	1.36 (3.50)	0.98 (2.59)	0.85 (2.34)
	None	0.68 (1.17)	0.65 (1.12)	1.16 (2.07)	1.26 (2.29)	0.97 (1.78)	1.03 (1.94)
Joint space narrowing score	Good	0.44 (1.11)	0.41 (1.17)	0.86 (2.27)	0.43 (0.84)	0.47 (1.23)	0.41 (1.16)
	Moderate	0.69 (1.80)	0.59 (1.43)	1.47 (2.86)	1.34 (2.76)	0.96 (2.25)	0.82 (1.96)
	None	0.56 (1.26)	1.00 (2.63)	1.13 (1.96)	1.35 (2.23)	0.91 (1.73)	1.22 (2.37)
Total score	Good	0.81 (1.93)	0.77 (2.09)	1.27 (3.31)	0.53 (0.84)	0.84 (2.06)	0.76 (2.04)
	Moderate	1.41 (3.03)	1.22 (2.60)	2.94 (6.30)	2.70 (6.03)	1.94 (4.49)	1.67 (4.01)
	None	1.25 (2.26)	1.65 (3.39)	2.29 (3.74)	2.61 (4.13)	1.88 (3.27)	2.25 (3.86)

Data are mean (SD).



**Figure 2** Improvements in physical function across European League Against Rheumatism (EULAR) states based on 28-joint Disease Activity Score (DAS28) C-reactive protein (CRP) and DAS28 erythrocyte sedimentation rate (ESR) for patients in the ATTAIn (Abatacept Trial in Treatment of Anti-TNF INadequate responders) and AIM (Abatacept in Inadequate responders to Methotrexate) trials (both treatment groups). Mean improvement from baseline in Health Assessment Questionnaire Disability Index (HAQ-DI), for combined abatacept and patients treated with placebo who were EULAR good, moderate or non-responders based on DAS28 (CRP) or DAS28 (ESR). A. Mean improvement from baseline to 6 months for patients in the ATTAIn trial. B. Mean improvement from baseline to 6 months for patients in the AIM trial. Error bars represent the standard error of the mean (SEM).

definitions, and would not be expected to result in more favourable results with one definition vs the other. Although the current analysis detected discrepancies in 17.6% of patients, these were moderate in magnitude; overall, there is general agreement between the two measures. The 2 DAS28 definitions also demonstrated agreement in classifying DAS28-defined remission, with the majority of discrepancies (33 cases) resulting in patients being classified as in remission with DAS28 (CRP) but not DAS28 (ESR). Considering the DAS28 (ESR) definition as the external comparison, there is support for the criterion validity of the DAS28 (CRP).

When a discrepancy occurs, the tendency is for the DAS28 (CRP) definition to yield a better response state. This has also been found in the work by Inoue *et al*<sup>23</sup> and Matsui *et al*.<sup>24</sup> To overcome this discrepancy, transforming the DAS28 (CRP) to more closely conform to the DAS28 (ESR) is an enticing idea. However, the concern is the generalisability of the transformation across different patient groups. Inoue *et al*<sup>23</sup> derived an adjustment factor based on regressing DAS28 (ESR) on DAS28 (CRP) (ie,  $DAS28 (ESR) = 1.01 DAS28 (CRP) + 0.590$ ). Applying this adjustment factor to the current dataset led to a larger percentage being classified with a worse response state using DAS28 (CRP) (2.9% classified as better and 12.9% classified worse, compared with the unadjusted results of 12.7% and 4.9%, respectively). Other approaches can be considered, for example, an adjustment factor could be based on regressing  $\ln(ESR)$  on  $\ln(CRP+1)$ , which is the essential difference between the two DAS28 definitions, and using this adjustment in the DAS28 (ESR) formula. Applying this adjustment, a more equitable division resulted, with 4.5% in a better and 8.6% in a worse state; again, the generalisability of the transformation may be an issue.

The construct validity of the DAS28 (CRP) was evaluated by examining the trends in radiographic and functional progression across the EULAR responder states based on the DAS28 (CRP). For both definitions of the DAS28 there was generally good agreement between reduced radiographic progression and EULAR responses, and classification as a good EULAR responder was predictive of a lower rate of radiographic progression in patients treated with abatacept. However, some deviations from this trend were observed. For the DAS28 (ESR) definition, radiographic scores were generally higher for the moderate responders than for non-responders in the placebo group. This was also true for the DAS28 (CRP) definition, in the abatacept and placebo groups. Changes in HAQ-DI scores showed very

**Table 4** Functional progression (Health Assessment Questionnaire Disability Index (HAQ-DI)) across the European League Against Rheumatism (EULAR) states based on the 28-joint Disease Activity Score (DAS28) C-reactive protein (CRP) and DAS28 erythrocyte sedimentation rate(ESR): mean change from baseline

Trial	EULAR response	Abatacept group		Placebo group		Combined group	
		CRP	ESR	CRP	ESR	CRP	ESR
ATTAIn (6 months)	Good	-1.00 (0.10)	-0.82 (0.13)	-0.56 (0.28)	-0.63 (0.38)	-0.95 (0.09)	-0.80 (0.12)
	Moderate	-0.49 (0.06)	-0.56 (0.06)	-0.21 (0.06)	-0.28 (0.06)	-0.43 (0.05)	-0.50 (0.05)
	None	-0.18 (0.06)	-0.21 (0.05)	-0.10 (0.06)	-0.07 (0.06)	-0.14 (0.04)	-0.14 (0.04)
AIM (6 months)	Good	-0.89 (0.06)	-0.97 (0.07)	-1.03 (0.16)	-1.20 (0.24)	-0.91 (0.05)	-0.99 (0.06)
	Moderate	-0.55 (0.04)	-0.57 (0.04)	-0.66 (0.06)	-0.71 (0.07)	-0.59 (0.03)	-0.61 (0.03)
	None	-0.11 (0.05)	-0.16 (0.06)	-0.18 (0.07)	-0.20 (0.07)	-0.15 (0.05)	-0.18 (0.05)

Data are mean (SD).

AIM, Abatacept in Inadequate responders to Methotrexate; ATTAIn, Abatacept Trial in Treatment of Anti-TNF INadequate responders.

similar linear trends for both definitions, with improvements in HAQ-DI across the EULAR responder states from none to moderate to good. These trends held for patients treated with abatacept and placebo, and demonstrate that improvements in the DAS28 (CRP) and DAS28 (ESR) are associated with improvements in physical function.

The level of sensitivity of the DAS28 (CRP) and DAS28 (ESR) is an important issue to consider, and differences could potentially contribute to discrepancies observed between the two measures. However, as the SRMs were identical for the two measures, and the relative efficiencies were comparable, suggesting the sensitivity to change of the DAS28 (CRP) and DAS28 (ESR) are very similar.

The analyses described here were performed only on patients for whom ESR and CRP measurements were available. The majority of patients who did not have both measurements had missing ESR data. This was due to technical issues associated with the fact that ESR assessments were performed locally on standard kits, whereas CRP was measured by a central laboratory, allowing electronic data transfer to the main database. Although only a subset of patients from each trial had CRP and ESR measurements, baseline demographics and clinical characteristics for this subset were comparable with those that did not have both measurements, suggesting that they were likely to be representative of the full patient population.

Two key findings have emerged from the current investigation. Firstly, the DAS28 (CRP) has been validated with respect to functional and radiographic progression, and the validation profile was similar to that based on ESR. Secondly, there was a tendency for the DAS28 (CRP) to yield a lower score and in some instances a better EULAR response (12.7% vs 4.9% for the CRP and ESR, respectively). Often the two definitions were generally comparable, and formulating a conclusion on the EULAR response state or remission status based on the DAS28 (CRP) definition provides a reasonable basis for assessing a patient, provided the user is aware of the tendency of the DAS28 (CRP) to yield a better response than the DAS28 (ESR). Currently, the same cut-off points for EULAR response states originally derived for the DAS28 (ESR) are also applied for the DAS28 (CRP). To increase the level of agreement between the two DAS28 formulations, one solution would be to derive a new set of cut-off points tailored for use with DAS28 (CRP). The development of a robust approach leading to set of cut-off points that can be generally applied across populations may prove difficult, as illustrated when the translation derived by Inoue *et al*<sup>23</sup> and Matsui *et al*<sup>24</sup> is applied to the current dataset. Exploration of a robust method that will translate DAS28 (CRP) to DAS28 (ESR) and the derivation of a corresponding set of cut-off points for the disease activity state for the CRP formulation are needed before a transformation of the DAS28 (CRP) definition is advocated. An appropriate research agenda encompassing the assembly of large and diverse datasets, and an analysis plan involving regression based models will help to provide further insight into the potential for achieving a better alignment of the DAS28 (CRP)-defined and the DAS28 (ESR)-defined EULAR response states.

In conclusion, we have provided a validation of the DAS28 (CRP) assessment and corresponding EULAR response states against radiographic progression and physical function. While the DAS28 (CRP) yielded a better EULAR response more often than the DAS28 (ESR), the validation profile was similar to that of the DAS28 (ESR), indicating that both definitions of the DAS28 criteria can be used as benchmarks to assess patient

improvement and treatment effect, and can aid in the description and interpretation of changes in disease activity in patients with RA.

**Acknowledgements:** The authors would like to thank Helen Clarke, Medicus International, for her editorial assistance.

**Funding:** This study is based upon clinical trial results from studies sponsored by Bristol-Myers Squibb, Princeton, New Jersey, USA. The authors had responsibility for the analysis and data interpretation, and all authors were involved in the drafting and critical revision of the article; in addition, all authors have seen and approved the final article for submission. This study was supported in part by an unrestricted research grant-in-aid from Bristol-Myers Squibb.

**Competing interests:** GW has received consultancies/speaking fees/honoraria from Bristol-Myers Squibb; J-CB is an employee of Bristol-Myers Squibb and has stock options; JT is an employee of Bristol-Myers Squibb and has stock options; MD has received consultancies/speaking fees/honoraria from Bristol-Myers Squibb, Abbott, Wyeth, Centocor and Schering Plough; MS has received consultancies/speaking fees/honoraria from Bristol-Myers Squibb and Centocor; JS has received consultancies/speaking fees/honoraria from Bristol-Myers Squibb; DA has received consultancies/speaking fees/honoraria from Bristol-Myers Squibb; PICMVR has received consultancies/speaking fees/honoraria from Bristol-Myers Squibb, Abbott, Wyeth, Novartis and Schering Plough.

**Ethics approval:** Ethics approval was obtained.

## REFERENCES

1. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, *et al*. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:727–35.
2. van der Heijde DM, van't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, *et al*. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;**49**:916–20.
3. 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.
4. Fransen J, van Riel PL. The Disease Activity Score and the EULAR response criteria. *Clin Exp Rheumatol* 2005;**23**:S93–9.
5. van Gestel AM, Haagsma CJ, van Riel PL. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;**41**:1845–50.
6. Fransen J, van Riel PL. DAS remission cut points. *Clin Exp Rheumatol* 2006;**24**:S029–32.
7. UMC Nijmegen. DAS score. <http://www.das-score.nl/> (accessed 12 April 2007).
8. 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. *Br J Rheumatol* 1996;**35**:1101–5.
9. Vander Cruyssen B, Van Looy S, Wyls B, Westhovens R, Durez P, Van den Bosch F, *et al*. DAS28 best reflects the physician's clinical judgment of response to infliximab therapy in rheumatoid arthritis patients: validation of the DAS28 score in patients under infliximab treatment. *Arthritis Res Ther* 2005;**7**:R1063–71.
10. Smolen JS, Han C, Bala M, Maini RN, Kalden JR, van der Heijde D, *et al*. Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study. *Arthritis Rheum* 2005;**52**:1020–30.
11. Weinblatt ME, Keystone EC, Furst DE, Kavanaugh AF, Chartash EK, Segurado OG. Long term efficacy and safety of adalimumab plus methotrexate in patients with rheumatoid arthritis: ARMADA 4 year extended study. *Ann Rheum Dis* 2006;**65**:753–9.
12. Fransen J, Welsing P, Keijzer Rd, Riel Pv. Disease Activity Scores using C-reactive protein: CRP may replace ESR in the assessment of RA disease activity. *Ann Rheum Dis* 2004;**62**:151.
13. Genovese MC, Becker JC, Schiff M, Luggen M, Sherrer Y, Kremer J, *et al*. Abatacept for rheumatoid arthritis refractory to tumor necrosis factor alpha inhibition. *N Engl J Med* 2005;**353**:1114–23.
14. Kremer JM, Genant HK, Moreland LW, Russell AS, Emery P, Abud-Mendoza C, *et al*. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2006;**144**:865–76.
15. Wells GA, Tugwell P, Kraag GR, Baker PR, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol* 1993;**20**:557–60.
16. Genant HK. Methods of assessing radiographic change in rheumatoid arthritis. *Am J Med* 1983;**75**:35–47.
17. Genant HK, Jiang Y, Peterfy C, Lu Y, Redei J, Countryman PJ. Assessment of rheumatoid arthritis using a modified scoring method on digitized and original radiographs. *Arthritis Rheum* 1998;**41**:1583–90.
18. Genant HK. Interleukin-1 receptor antagonist treatment of rheumatoid arthritis patients: radiologic progression and correlation of Genant/Sharp and Larsen scoring methods. *Semin Arthritis Rheum* 2001;**30**:26–32.

19. **Kremer J**, Genant H, Moreland L, Russell A, Emery P, Abud-Mendoza C, *et al*. Effects of abatacept in patients with active rheumatoid arthritis despite methotrexate: a randomized trial. *Ann Intern Med* 2006;**144**:865–76.
20. **Bland JM**, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**1**:307–10.
21. **Wells G**, Li T, Maxwell L, Maclean R, Tugwell P. Responsiveness of patient reported outcomes including fatigue, sleep quality, activity limitation, and quality of life following treatment with abatacept for rheumatoid arthritis. *Ann Rheum Dis* 2008;**67**:260–5.
22. **Kushner I**. C-reactive protein in rheumatology. *Arthritis Rheum* 1991;**34**:1065–8.
23. **Inoue E**, Yamanaka H, Hara M, Tomatsu T, Kamatani N. Comparison of Disease Activity Score (DAS)28-erythrocyte sedimentation rate and DAS28-C-reactive protein threshold values. *Ann Rheum Dis* 2007;**66**:407–9.
24. **Matsui T**, Kuga Y, Kaneko A, Nishino J, Eto Y, Chiba N, *et al*. Disease Activity Score 28 (DAS28) using CRP underestimates the disease activity and overestimates the EULAR response criteria compared with DAS28 using ESR in a large observational cohort of rheumatoid arthritis patients in Japan. *Ann Rheum Dis* 2007;**66**:1221–6.