

EXTENDED REPORT

Definition of treatment response in rheumatoid arthritis based on the simplified and the clinical disease activity index

Daniel Aletaha,¹ Jose Martinez-Avila,¹ Tore K Kvien,² Josef S Smolen^{1,3}

► An additional table is published online only. To view this file please visit the journal online (<http://ard.bmj.com/content/early/recent>).

¹Division of Rheumatology, Medical University of Vienna, Vienna, Austria

²Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway

³Second Department of Medicine, Hietzing Hospital, Vienna, Austria

Correspondence to

Dr Daniel Aletaha, Medical University of Vienna, Department of Rheumatology, Spitalgasse 23, 1090 Vienna, Austria; daniel.aletaha@meduniwien.ac.at

Handling editor

Johannes WJ Bijlsma

Received 4 July 2011

Accepted 8 February 2012

ABSTRACT

Background The simplified disease activity index (SDAI) and the clinical disease activity index (CDAI) are established instruments to measure disease activity in rheumatoid arthritis (RA). To date, no validated response definitions for the SDAI and CDAI are available.

Objective The authors aimed to define minor, moderate and major response criteria for the SDAI.

Methods The authors used data from two clinical trials on infliximab versus methotrexate in early (ASPIRE) or established (ATTRACT) RA, and identified the three SDAI cutpoints based on the best agreement (by κ statistics) with the American College of Rheumatology (ACR)20/50/70 responses. Cutpoints were then tested for different aspects of validity in the trial datasets and in a Norwegian disease modifying antirheumatic drug prescription dataset (NOR-DMARD).

Results Based on agreement with the ACR response, the minor, moderate and major responses were identified as SDAI 50%, 70% and 85% improvement. These cutpoints had good face validity concerning the disease activity states achieved by the different response definitions. Construct validity was shown by a clear association of increasing SDAI response categories with increasing levels of functional improvement, achievement of better functional states and lower annual radiographic progression. Across SDAI 50/70/85, the sensitivities regarding a patient-perceived improvement decreased (73%/39%/22%) and the specificities increased (61%/89%/96%) in a meaningful way. Further, the cutpoints discriminated the different treatment arms in ASPIRE and ATTRACT. The same cutpoints were used for the CDAI, with similar results in the validation analyses.

Conclusion These new response criteria expand the usefulness of the SDAI and CDAI for their use as endpoints in clinical trials beyond the definition of disease activity categories.

INTRODUCTION

Various means are available to assess disease activity of rheumatoid arthritis (RA). Since the disease per se is multi-faceted,¹ typically composite disease activity measures are used that combine different variables.²⁻⁴ The most commonly used indices are the disease activity score based on a 28-joint count (DAS28), the simplified disease activity index (SDAI) and the clinical disease activity index (CDAI), all of which can be applied in clinical trials and clinical practice.⁵⁻¹⁰ The use of these indices to follow disease activity over time has also become a

very important aspect in the care for patients with RA.¹¹⁻¹⁶

Response criteria inherently require assessments at least at two time points. The most commonly used response criteria in clinical trials are the ones of the American College of Rheumatology (ACR),⁸ which have been derived based on the discrimination of responses between active drug and placebo. The most recently proposed revision of the ACR criteria has maintained its focus on response.¹⁷ The response criteria of the European League Against Rheumatism (EULAR) are based on the DAS28 and were developed on the basis of the notion that not only the change in disease activity upon therapeutic intervention was important, but, potentially also the disease activity state that has been reached.¹⁸

Although we have recently shown that it is clinically more important to achieve a good state than merely a good response,¹⁹ it has been recommended to maintain the reporting of response rates in addition to reporting of disease activity states, especially for the short-term perspective of clinical trials.^{20 21}

The newer SDAI and CDAI currently have no validated response measures, but in a previous study we have shown that the patients' perception of response is dependent on the starting level of disease activity.¹⁹ Using the ACR response definition as reference, we aimed to take this observation further by developing and validating response criteria for these two indices that (A) are simple, (B) relate to patients' perception of a response, (C) are based on relative changes from a baseline (to reflect the aforementioned findings), and (D) predict important outcomes of RA.

METHODS

Overall study design

Derivation of response cutpoints

Our study design is based on four a priori assumptions: (A) *Response metric*: on the basis of the results from an earlier study,¹⁹ we defined the response metric of the SDAI to be based on percentage change in score, which was best correlated with the patients' perception of improvement; (B) *Gold standard*: we will define three levels of cutpoints based on the best agreement with the traditional ACR20/50/70 system, which is the most established way of determining response in clinical trials of RA; these response levels will be labelled as 'minor', 'moderate' and 'major'; (C) *Simplicity*: to maintain the simplicity inherent to the concept of the SDAI and CDAI, we will test and accept

Table 1 Baseline characteristics for the different cohorts used (mean±SD, unless indicated otherwise)

Variable	NOR-DMARD	ASPIRE (DMARD naïve)		ATTRACT (background MTX)	
		MTX	MTX + IFX*	PLC	IFX*
Patient (n)	1129	242	579	68	262
Age (years)	55.3±26.9	49.6±12.7	49.5±12.8	50.8±11.5	52.7±11.4
Female (%)	73.3	74.2	70.3	81.3	76.9
Rheumatoid factor positivity (%)	67.8	58.5	51.7	77.3	70.5
Disease activity score 28	4.97±1.36	6.20±1.02	6.17±1.04	6.55±1.02	6.7±1.0
Duration of RA (years)	8.2±9.6	0.8±0.7	0.8±0.7	10.4±8.3	10.2±8.0
Health assessment questionnaire (0-3)	0.7±0.5†	1.5±0.6	1.5±0.7	1.7±0.6	1.7±0.6
Simplified disease activity index	26.9±13.9	38.1±11.7	38.2±12.2	45.1±14.7	46.9±14.7
Clinical disease activity index	24.6±13.1	35.5±10.5	35.2±11.1	41.7±13.8	43.2±13.4
Swollen joint count (0-28)	7.1±5.6	8.9±4.2	9.2±4.4	15.0±6.3	14.9±6.3
Tender joint count (0-28)	8.2±6.8	14.2±6.2	13.4±6.3	14.6±7.7	16.2±7.3
C-reactive protein (mg/dl)	2.26±2.77	2.54±2.95	2.90±3.31	3.81±3.69	3.70±3.78
Erythrocyte sedimentation rate (mm/h)	28.7±22.7	44.4±27.2	45.6±29.2	48.2±23.7	49.4±24.0
Patient global assessment (mm)	52.2±24.1	60.4±25.1	61.1±24.7	59.2±24.9	61.1±22.7
Evaluator global assessment (mm)	40.6±19.0	65.2±17.2	65.8±18.6	62.2±17.4	60.0±17.9
Radiographic score	-	11.4±16.2	11.1±14.8	80.9±79.4	69.6±64.7

*The two infliximab doses from ASPIRE were pooled, and the four different doses (or dosing intervals) from ATTRACT were also pooled.

†Physical disability measured by the modified Health assessment.

NOR-DMARD, Norwegian disease modifying antirheumatic drug prescription dataset; IFX, infliximab; MTX, methotrexate; RA, rheumatoid arthritis; PLC, placebo.

response cutpoints between 5% and 95% in steps of 5; (D) *Similarity*: given that SDAI and CDAI are identical except for the inclusion of C-reactive protein in the SDAI, we will aim to develop definitions of response that can be applied to both indices in the same way; therefore, only SDAI will be used for derivation of response cutpoints, but validations of these cutpoints will be performed for both scores.

Validation of response cutpoints

We will perform fourfold testing of the derived cutpoints for (A) *face validity*, by investigating which states of disease activity are achieved by patients when they show a certain level of response; (B) *construct validity*, by investigating to what extent the different response levels are associated with important outcomes of the disease, namely structural and functional outcomes; (C) *consistency with patient perception*, by analysing the sensitivities and specificities which the defined responses exhibit concerning the patients' perception of improvement; and (D) *discriminant validity*: by studying how the different response levels discriminate responses achieved with an active agent compared with a placebo or comparator drug.

Patients and datasets

We used three data sources, for the purpose of our study: two from clinical trials and one observational dataset. One trial was on established RA, the anti-TNF trial in rheumatoid arthritis with concomitant therapy (ATTRACT),²² comparing infliximab (IFX) plus methotrexate (MTX) versus placebo and methotrexate in patients with active disease despite MTX therapy and the other one on early RA, the active controlled study of patients receiving infliximab for treatment of rheumatoid arthritis of early onset (ASPIRE),²³ comparing IFX plus MTX versus MTX alone in MTX-naïve patients with RA. For both trials, we had 1 year data on courses of disease activity, physical function and radiographic outcomes (total Sharp score) available. An 80% random patient sample of both trials was kindly provided by Centocor (Malvern, PA, USA). For all analyses, we pooled the different dosing groups of IFX (two in ASPIRE; four in ATTRACT).

The source of the observational data was a Norwegian prescription dataset (the NOR-DMARD study).²⁴ We identified and analysed the first documented course with any disease-modifying

antirheumatic drug (DMARD) in each patient. Available data included measures of disease activity and function (modified health assessment questionnaire, mHAQ) at start of the respective DMARD therapy (baseline) and after 3, 6 and 12 months follow up; radiographic data were not available.

NOR-DMARD allowed us to investigate the consistency with patient perception (see validation of response cutpoints), as it includes a patient-reported measure of perceived improvement on DMARD therapy¹⁹. At all visits except the baseline visit, patients were asked to assess the improvement of their disease activity on a 5-point Likert scale. The wording of the question was: 'Since you started treatment in this follow-up study, has your rheumatic disease improved, been unchanged or become worse?' (originally in Norwegian language: 'Siden du startet behandling i denne oppfølgingsundersøkelsen, er du blitt bedre, uforandret eller verre i din revmatiske sykdom?'), and the response options were: 'considerably better', 'better', 'unchanged', 'worse', and 'considerably worse'.

In all three datasets, we calculated the composite disease activity measures, SDAI⁹ and CDAI,¹⁰ and the DAS28 from the original data.

Derivation of the best response cutpoints

We derived three response cutpoints for the SDAI based on the point (per cent change of SDAI) of best statistical agreement (Cohen's κ) with ACR20 (minor response), ACR50 (moderate response) or ACR70 (major response) at 6 months. The κ value ranges from 0 to 1, with a value of 0 indicating agreement as expected by chance, and one indicating perfect agreement. Based on this analysis in the three active treatment arms of the trials (ASPIRE: MTX+IFX and MTX alone; ATTRACT: MTX+IFX), we obtained and summarised the best cutpoints for the three '% SDAI change' response levels to be further evaluated for different aspects of validity.

Testing and validation of the derived response cutpoints

Clinical face validity

We first identified which states of disease activity would be reached by patients who achieved a response by the minor, moderate and major response cutpoints of the SDAI using

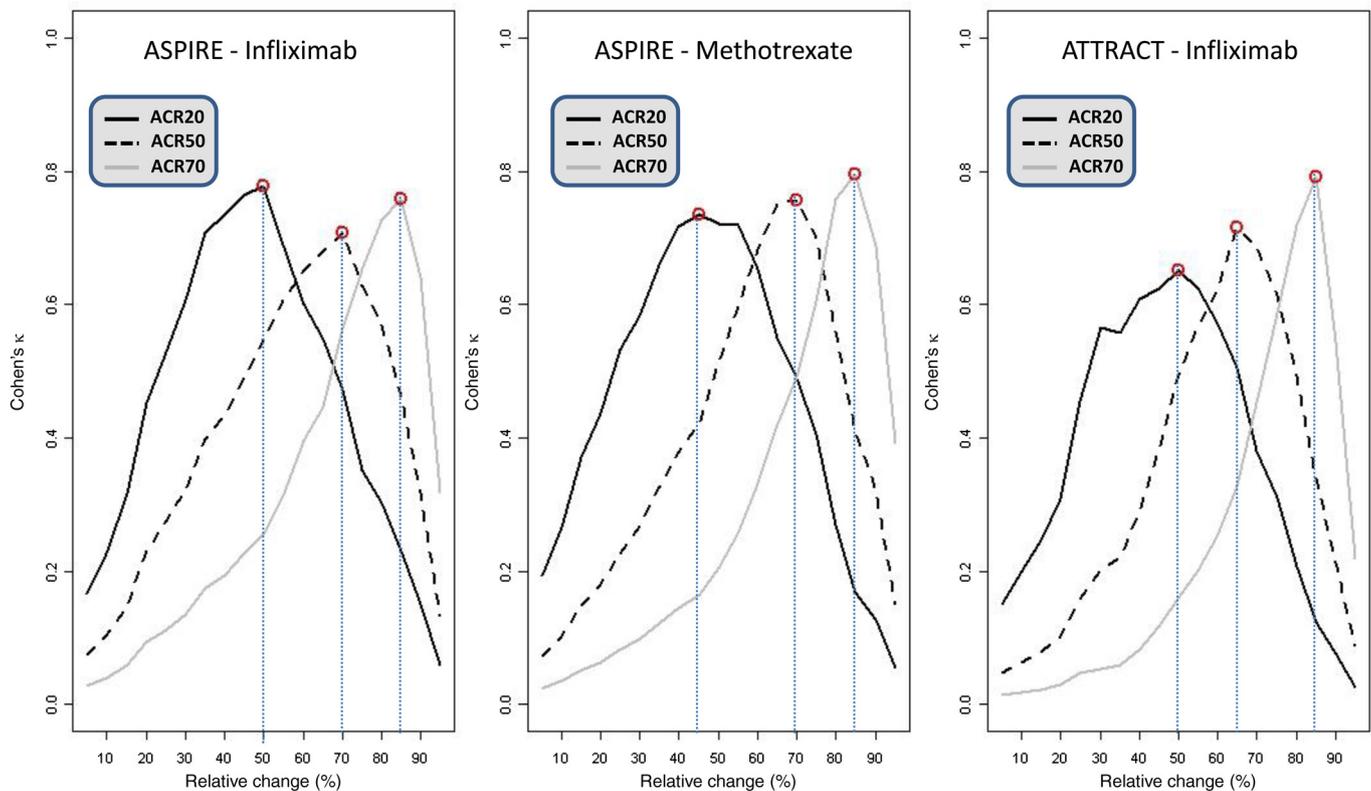


Figure 1 Identification of SDAI cutpoints. Lines indicate the level of agreement of SDAI cutpoints (X-axis) with ACR20 response (solid black lines, corresponding to SDAI minor response), ACR50 responses (broken black lines, corresponding to SDAI moderate response) and ACR70 (solid grey lines; corresponding to SDAI major response). The individual panels depict the different trial arms analysed: ASPIRE infliximab (left panel), ASPIRE methotrexate (middle panel) and ATTRACT infliximab (right panel). κ Values of 0.5 indicate agreement as expected just by chance, while κ values of 1 indicate perfect agreement. The vertical lines indicate the points of maximum agreement, which were used to identify the best SDAI cutpoints. Analyses are based on the 6 months' data. ACR, American College of Rheumatology; SDAI, simplified disease activity index.

the 6-month time point as a clinically relevant time to assess response to therapy. Face validity would suggest that patients with major response should mostly be reaching low disease activity or remission, and that this proportion would decrease when the cutpoints for moderate response or minor response are evaluated. The 6-month cutpoint was chosen because recent therapeutic strategies suggest that a good clinical response should be reached within 6 months from start of therapy.¹⁶ Also, we tested the consistency of the SDAI response levels in comparison with the EULAR response criteria (which are based on DAS28 response and state achieved).²⁵

Construct validity

A construct of the disease, RA, is the fact that (A) higher levels of disease activity are associated with higher levels of functional impairment, and thus (B) improvement of disease activity is reflected also in an improvement of physical function, and that (C) ongoing disease activity leads to joint destruction, although for the latter it has been shown that the disease activity state reached, as opposed to the response level reached, is important.²⁶

We tested these outcomes in all arms of the ATTRACT and ASPIRE trials and in NOR-DMARD; in the latter, we did not examine joint damage as there were no radiographic data available. We looked at both, the 6-month and the 12-month time point of responses and outcomes. For radiographic analysis we compared the 1-year radiographic progression between SDAI non-responders, SDAI minor (but not moderate) responders, SDAI moderate (but not major) responders, and SDAI major responders at 6 and 12 months.

Validity regarding patient perspective

We identified sensitivities and specificities for the three response levels regarding the degree of improvement as reported by the patients on a 5-point Likert scale after 6 and 12 months (see above).¹⁹ For the analysis of the patients' relevant response, the status of the anchor was defined as follows: 'considerably worse' or 'worse' or 'unchanged' = no subjective improvement; 'improved' or 'considerably improved' = subjective perception of improvement. Since in clinical practice a patient with worsening SDAI would not even be considered to be assessed for the presence or absence of a relevant response, we excluded all patients with negative relative SDAI changes. We used a receiver operating characteristics (ROC) curve analysis to show the changes in sensitivities and specificities across the range of different cutpoints.²⁷

Discriminative validity

Another validity criterion for a response measure in RA is its ability to discriminate active treatment from control treatment in clinical trials. We used the ATTRACT and ASPIRE trials for this assessment. In ATTRACT, infliximab was compared with true placebo (both added to the background of insufficiently effective methotrexate). In ASPIRE, IFX plus MTX was compared with MTX alone in a MTX-naïve setting. We used χ^2 test to compare the statistical power to discriminate the proportion of responders between active and comparator regimens in the two trials at 1 year.

For all our analysis we used the R software.²⁸

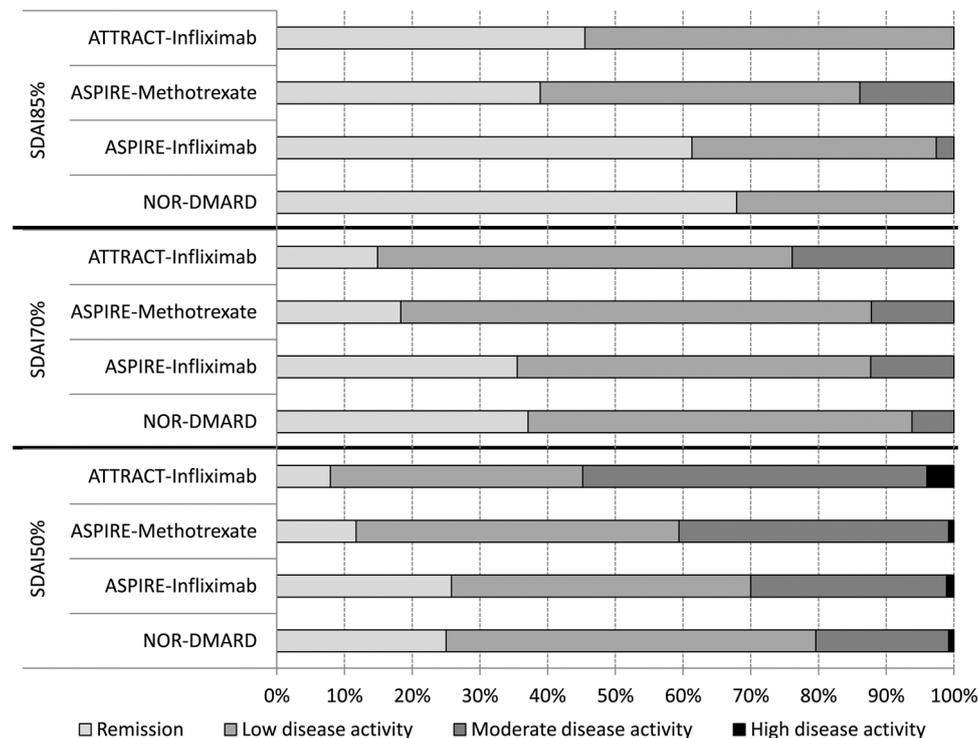


Figure 2 Face validity of response definitions. Face validity in this context relates to reaching a sensible disease activity state for patients who have achieved a particular response. The horizontal bars indicate the proportion of patients who achieved the different disease activity states by having a SDAI 50, 70 or 85 response. We analysed the active arms of ASPIRE (infliximab, methotrexate) and ATTRACT (infliximab) as well as patients newly prescribed DMARDs in NOR-DMARD. A minor response of SDAI allowed almost all patients to move out of the high disease activity state; when a moderate SDAI response was achieved, then low disease activity was the most frequent disease activity state; and when a major response was reached, then remission was the most common disease activity state. Data are based on the 6 months' evaluations. NOR-DMARD, Norwegian disease modifying antirheumatic drug prescription dataset; SDAI, simplified disease activity index.

RESULTS

Patients

Table 1 shows the baseline characteristics of patients identified in NOR-DMARD and in the two arms each of the ASPIRE and the ATTRACT trials. From NOR-DMARD we included all DMARD therapies that were initiated, mostly methotrexate (37.2%) and biological DMARDs (40.8%).

Identification of the best response cutpoints

Figure 1 displays the κ values for increasing levels of relative SDAI responses (% SDAI change) compared with ACR20/50/70 responses. It can be seen that the optimal cutpoints for SDAI minor response (ie, the one best corresponding to ACR20 responses) were 50%, 45% and 50% changes in SDAI in the three active treatment arms of the trials, respectively. The best cutpoints for SDAI moderate response (corresponding to ACR50) were 70%, 70% and 65%; for SDAI major response (corresponding to ACR70) they were 85% in all three arms.

Based on these results, the SDAI minor response level was defined as a relative change of SDAI of 50%, SDAI moderate as a relative change of 70% and SDAI major response as a change in SDAI of 85%. The agreement with ACR response levels was very similar when the CDAI responses were tested (data not shown), and we therefore proceeded to identical cutpoints for SDAI and CDAI, as planned.

In a sensitivity analysis, we evaluated all patients for whom data were available, including those with negative SDAI changes, and the total area under the ROC curve was 0.732,

that is, virtually identical with the 0.752 that we obtained after excluding patients with negative SDAI changes.

We next tested the new response definitions in several different ways.

Clinical face validity

Figure 2 depicts the disease activity states that patients with a minor, moderate or major response have reached at 6 months, using the three active arms of ASPIRE and ATTRACT. Achieving a minor (SDAI50) response led to the virtual absence of high disease activity, while in moderate (SDAI70) responders low disease activity was the dominating state outcome. In major (SDAI85) responders, remission rates reached approximately 40%–60% and very few patients remained in moderate disease activity.

There was reasonable consistency of the new SDAI criteria in comparison with the EULAR response criteria (online supplementary table S1).

Construct validity

Tables 2 and 3 summarise the results of the comparisons of HAQ changes, endpoint HAQ values and 1 year radiographic progression with 6 month SDAI response status for the different cohorts used. It can be seen that functional responses and the HAQ scores reached were associated with the level of SDAI response (table 2). As expected, functional states that were reached even with SDAI85 response in the ATTRACT trial were worse than functional states seen in the ASPIRE trial. This relates to the greater damage-related irreversible functional disability in long-standing RA that remains unaffected

Table 2 Functional construct validity of SDAI response cutpoints (6 months' data; all treatment arms, infliximab + methotrexate and controls, were pooled for the two trials, ASPIRE and ATTRACT; mean±SD). $p < 0.001$ for global test in all cohorts and all analyses

SDAI response	HAQ improvement (change)			HAQ reached (status)		
	ASPIRE	ATTRACT	NOR-DMARD*	ASPIRE	ATTRACT	NOR-DMARD*
No response	-0.43±0.54	-0.25±0.53	-0.11±0.40	1.11±0.63	1.50±0.65	0.53±0.42
SDAI50	-0.75±0.57	-0.51±0.54	-0.28±0.38	0.76±0.60	1.16±0.67	0.35±0.35
SDAI70	-0.95±0.57	-0.76±0.49	-0.40±0.45	0.58±0.55	0.70±0.58	0.30±0.33
SDAI85	-1.28±0.60	-0.73±0.40	-0.55±0.54	0.20±0.32	0.70±0.64	0.16±0.30

*Physical disability measured by mHAQ.

mHAQ, modified Health assessment questionnaire; NOR-DMARD, Norwegian disease modifying antirheumatic drug prescription dataset; SDAI, simplified disease activity index;

Table 3 Structural construct validity of SDAI response cutpoints in the active treatment arms of ASPIRE and ATTRACT (response status at 6 month, radiographic progression over year 1; mean±SD)

SDAI response category	ASPIRE		ATTRACT
	MTX	MTX+IFX	IFX
No response	5.6±12.6 (n=59)	1.8±6.4 (n=107)	2.9±8.5 (n=50)
SDAI50	3.9±9.8 (n=42)	0.6±4.5 (n=79)	0.3±3.5 (n=61)
SDAI70	2.2±6.6 (n=35)	0.3±4.3 (n=98)	0.1±5.5 (n=49)
SDAI85	-0.3±5.9 (n=44)	-0.4±5.2 (n=179)	0.3±4.5 (n=43)
	p=0.09	p=0.01	p=0.67

IFX, infliximab; MTX, methotrexate; SDAI, Simplified Disease Activity Indices.

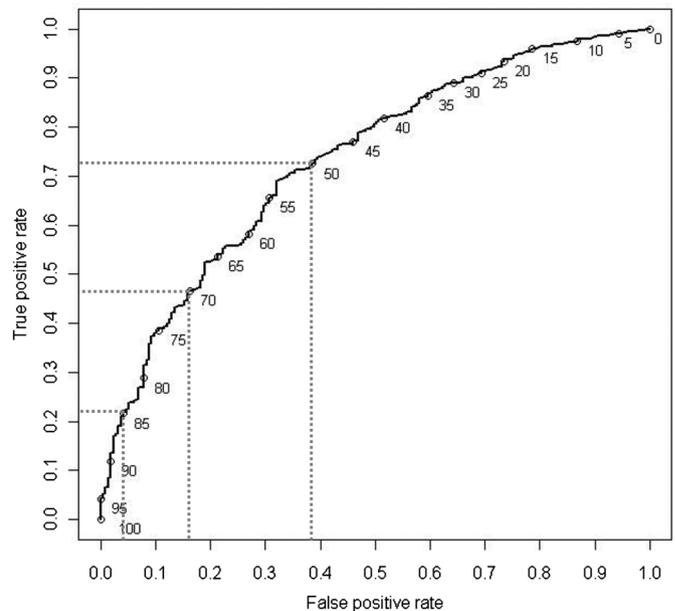
by improving disease activity.²⁹ Average radiographic progression was much lower in patients with greater levels of response by the SDAI in both the ASPIRE and the ATTRACT trials (table 3). When progression of joint damage was compared across patients with different ACR response levels (ACR non-responders, ACR20, ACR50 and ACR70), the association was similar (data not shown). The lack of significant effects on radiographic progression in the SDAI85 responders of ATTRACT was also seen with the ACR70 responders. This needs to be viewed in the context of the smaller sample size and of the completers analysis used in the present study, compared with the original trial.

Validity regarding patient perspective

The proportion of patients in NOR-DMARD who rated themselves much worse, worse, unchanged, improved and much improved were 0.5%, 6.6%, 19.2%, 43.7% and 30%, respectively, at 6 months, and 0.8%, 9.8%, 21.3%, 39.4% and 28.7%, respectively, at 12 months.

The expectation would be that the minor response cutpoint is sensitive in identifying patients who consider themselves improved (ie, be inclusive, few false negatives), while the major response should be specific by *only* identifying those who also subjectively feel improved (ie, be strict; few false positives). The sensitivities and specificities of relative changes of SDAI for patients' perception of improvement (improved/much improved vs rest) are shown in the ROC curve depicted in figure 3. The minor (SDAI50) response cutpoint showed a good sensitivity of 72.5% and specificity of 61.4% regarding the patients' perception of improvement. The moderate (SDAI70) cutpoint showed a moderate sensitivity of 46.6% paired with a high specificity of 83.7%. Finally, the major (SDAI85) cutpoint identified patient-reported responses with a low sensitivity of 21.6% but an excellent specificity of 95.8%.

Thus, three out of four patients with a subjective response are identified with the minor response definition. Vice versa, the

**Figure 3** Consistency with patient perception. Receiver operating characteristics (ROC) curve showing the rate of false positives (1-specificity) and true positives (sensitivity) for each increment of the test variable 'relative SDAI change at 1 year'. For better assessment of where cutpoints lie, the per cent change in SDAI or CDAI are marked on the ROC curve in steps of 5%. Data are based on response evaluation at 6 months. CDAI, clinical disease activity index; SDAI, simplified disease activity index.

major cutpoint is highly specific as it identifies only 4% patients who would not consider themselves improved.

Discriminative validity

Table 4 shows the response rates in the different arms of the ASPIRE and the ATTRACT trials, and provides the p values resulting from the comparison of response rates between the active and the comparator arm in each trial when the different response levels are used. In the ASPIRE trial of early RA, the SDAI minor, moderate and major response levels discriminated response to infliximab plus MTX compared with active MTX at p values of 1.4×10^{-3} , 3.4×10^{-4} and 5.2×10^{-5} , respectively. In the ATTRACT trial of established RA, infliximab was discriminated from placebo therapy (both with background MTX) at p values of 2.3×10^{-4} , 7.9×10^{-3} and 4.2×10^{-1} , respectively. The somewhat lower p values than in the original publication have to be viewed in the context of the smaller sample size and the completers analysis used in the present study compared with the original trial.

Table 4 Discriminative validity in the ASPIRE and ATTRACT trials at 6 months.

SDAI response category	ASPIRE			ATTRACT		
	MTX	MTX+IFX	p Value	PLC	IFX	p Value
SDAI50	65.2%	77.8%	1.4×10^{-3}	25.7%	60.6%	2.3×10^{-4}
SDAI70	39.8%	55.8%	3.3×10^{-4}	10.0%	32.2%	7.9×10^{-3}
SDAI85	14.4%	30.2%	5.2×10^{-5}	5.5%	10.6%	4.2×10^{-1}
ACR20	59.9%	73.4%	4.8×10^{-4}	40.9%	60.0%	2.9×10^{-2}
ACR50	36.5%	55.7%	5.5×10^{-6}	13.6%	32.3%	2.1×10^{-2}
ACR70	16.3%	32.5%	2.1×10^{-5}	4.5%	12.2%	2.3×10^{-1}

ACR, American College of Rheumatology; IFX, infliximab; MTX, methotrexate; SDAI, simplified disease activity index.

Results for the CDAI

Using the same cutpoints as for the SDAI, the four response categories of SDAI and CDAI showed an excellent agreement: looking at the ASPIRE trial, the ATTRACT trial and the NOR-DMARD dataset, each at the 6- and 12-months time point, the weighted κ values ranged between 0.94 and 0.98. However, We repeated all validation analyses subsequently, as shown above for the SDAI, also for the CDAI, which, as expected, led to the same conclusions (data not shown).

DISCUSSION

Many instruments are currently available for measuring disease activity in RA. These include, typically, the DAS28, but more recently also the SDAI, which had been introduced to provide a score that is simple to calculate and avoids weighting and transformation. Finally, the CDAI was developed to make formal disease activity assessment possible in situations when acute phase measures were not available. For all these indices, disease activity states have been defined.³⁰ While the DAS28 is the basis for the EULAR response criteria,³¹ which require change and status achievement, no formal criteria are available to evaluate achievement of a response for the SDAI and the CDAI.

In the present study, we newly introduce response criteria for the SDAI and CDAI. They are graded into mild, moderate and major responses, which will allow their application depending on the clinical question, or the specific purpose of a clinical trial. Both indices have recently been used in major clinical trials,^{32,33} and the new ACR/EULAR remission criteria also use a SDAI-based definition of remission.^{34,35} Moreover, current recommendations on data reporting in clinical trials suggest to present both disease activity states and responses.²⁰ Therefore, the elaboration of SDAI- and CDAI-based definitions of response is certainly overdue.

The anchor for determining response levels for the SDAI was the currently most established response classification system, namely the ACR improvement definition.⁸ This resulted in response requirements of 50% for mild, 70% for moderate and 85% for major improvement. In the datasets analysed, the minor SDAI response excluded almost all patients with a high disease activity state, while patients fulfilling the major SDAI response criteria were almost all in low disease activity or remission, an observation that lends a great deal of face validity to these response and improvement levels.

Of note, in contrast to the SDAI, the calculation of response in the ACR definition is based on a Boolean approach. In addition, the ACR response definition requires improvement in both swollen and tender joints, while in the SDAI, as a summative index, smaller improvements in one can be compensated by larger improvements in the other. All this may explain why

higher relative changes of SDAI corresponded to lower levels in the ACR (ACR20→SDAI50; ACR50→SDAI70; and ACR70→SDAI85). The similar proportions and similar outcomes in the validation analysis between corresponding ACR and SDAI response levels (as partly shown in the results section), further support the notion that the differences in cutpoints are merely related to conceptual differences in calculation of a response. It has to be emphasised that similar to the ACR responses, SDAI responses are inclusive, that is, an SDAI50 response also always comprises an SDAI70 and SDAI85 response.

A strength of the present study is that the newly derived criteria were analysed and validated in several cohorts. Each of the datasets provided a unique aspect to the validation, with ASPIRE providing trial data on early RA patients and ATTRACT on established RA, both with the possibility to also investigate structural outcomes. Finally, the NOR-DMARD dataset allowed to investigate the concordance of response definitions with the perception of improvement in a real-life cohort.

A limitation of a secondary data analysis like the present study is that the data have not been prospectively collected for the purpose of the study. Furthermore, it has to be borne in mind that the efficacy data from the trials (eg, radiographic progression or functional data) may greatly differ compared with the original reports of these trials. This is partly due to the 80% data cut, but particularly related to the fact that here data were analysed 'as observed', and not imputed or carried forward as in the main reports.

The aim of our analysis was to derive response criteria, which not only comprise all aspects of validity but also conform to the conceptual background of the SDAI and CDAI, namely simplicity and ease of use. Although these response definitions may still require the use of a calculator, the mode of calculation is much simpler than the Boolean-based integration of individual response calculations, as needed in the ACR response evaluation or other response assessments.

In summary, the new response criteria for the SDAI and CDAI conform to all aspects of validity, face, construct, correlational and discriminative validity. Moreover, they pay tribute to prior findings that patients judge improvement on therapies in a relative manner comparing it with where they came from when treatment was initiated (the 'baseline'). Thus, achievement of a relative change by 50%, 70% or 85% can be easily calculated, can be used to summarise efficacy data for a group of patients across all levels of baseline disease activity and, therefore, constitutes an easily useable tool. These definitions expand the convenience and reliability of the SDAI and CDAI to the level of response assessment.

Contributors DA and JSS contributed to the study design. DA, JM-A and JSS analysed the data. TKK provided the data and along with DA and JSS drafted the manuscript.

Funding This study was supported through Coordination Theme 1 (Health) of the European Community's FP7; Grant Agreement number HEALTH-F2-2008-223404 (Masterswitch). This is a publication of the Joint and Bone Center for Diagnosis, Research and Therapy of Musculoskeletal Disorders of the Medical University of Vienna.

Competing interests DA was involved in the development of SDAI and CDAI, received consulting and speaking honoraria from MSD. JM-A and T KK declare no competing interests. JSS was involved in development of SDAI and CDAI and received consulting and speaking honoraria from Centocor and MSD.

Acknowledgements The authors thank Centocor for providing the random sample of the original data from the ASPIRE and ATTRACT trials, and Dr Daniel Baker (Centocor) for the thoughtful discussions.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- Felson DT, Anderson JJ, Boers M, *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.
- Goldsmith CH, Smythe HA, Helewa A. Interpretation and power of a pooled index. *J Rheumatol* 1993;**20**:575–8.
- Bombardier C, Tugwell P. A methodological framework to develop and select indices for clinical trials: statistical and judgmental approaches. *J Rheumatol* 1982;**9**:753–7.
- van der Heijde DM, van't Hof MA, van Riel PL, *et al*. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. *Ann Rheum Dis* 1992;**51**:177–81.
- Paulus HE, Egger MJ, Ward JR, *et al*. Analysis of improvement in individual rheumatoid arthritis patients treated with disease-modifying antirheumatic drugs, based on the findings in patients treated with placebo. The Cooperative Systematic Studies of Rheumatic Diseases Group. *Arthritis Rheum* 1990;**33**:477–84.
- van der Heijde DM, van't Hof MA, van Riel PL, *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.
- Prevoo ML, van't Hof MA, Kuper HH, *et al*. 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.
- Felson DT, Anderson JJ, Boers M, *et al*. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:727–35.
- Smolen JS, Breedveld FC, Schiff MH, *et al*. A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology (Oxford)* 2003;**42**:244–57.
- Aletaha D, Nell VP, Stamm T, *et al*. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. *Arthritis Res Ther* 2005;**7**:R796–806.
- Grigor C, Capell H, Stirling A, *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.
- Smolen JS, Sokka T, Pincus T, *et al*. A proposed treatment algorithm for rheumatoid arthritis: aggressive therapy, methotrexate, and quantitative measures. *Clin Exp Rheumatol* 2003;**21**(5 Suppl 31):S209–10.
- Goekoop-Ruiterman YP, de Vries-Bouwstra JK, Allaart CF, *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.
- Aletaha D, Funovits J, Keystone EC, *et al*. Disease activity early in the course of treatment predicts response to therapy after one year in rheumatoid arthritis patients. *Arthritis Rheum* 2007;**56**:3226–35.
- Smolen JS, Aletaha D, Bijlsma JW, *et al*. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;**69**:631–7.
- Smolen JS, Landewé R, Breedveld FC, *et al*. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs. *Ann Rheum Dis* 2010;**69**:964–75.
- A proposed revision to the ACR20: the hybrid measure of American College of Rheumatology response. *Arthritis Rheum* 2007;**57**:193–202.
- van Gestel AM, Anderson JJ, van Riel PL, *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.
- Aletaha D, Funovits J, Ward MM, *et al*. Perception of improvement in patients with rheumatoid arthritis varies with disease activity levels at baseline. *Arthritis Rheum* 2009;**61**:313–20.
- Aletaha D, Landewe R, Karonitsch T, *et al*. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Arthritis Rheum* 2008;**59**:1371–7.
- Aletaha D, Landewe R, Karonitsch T, *et al*. Reporting disease activity in clinical trials of patients with rheumatoid arthritis: EULAR/ACR collaborative recommendations. *Ann Rheum Dis* 2008;**67**:1360–4.
- Lipsky PE, van der Heijde DM, St Clair EW, *et al*. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. *N Engl J Med* 2000;**343**:1594–602.
- St Clair EW, van der Heijde DM, Smolen JS, *et al*. Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. *Arthritis Rheum* 2004;**50**:3432–43.
- Kvien TK, Heiberg, Lie E, *et al*. A Norwegian DMARD register: prescriptions of DMARDs and biological agents to patients with inflammatory rheumatic diseases. *Clin Exp Rheumatol* 2005;**23**(5 Suppl 39):S188–94.
- van Gestel AM, Prevoo ML, van't Hof MA, *et al*. 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.
- Aletaha D, Funovits J, Smolen JS. The importance of reporting disease activity states in rheumatoid arthritis clinical trials. *Arthritis Rheum* 2008;**58**:2622–31.
- Sing T, Sander O, Beerenwinkel N, *et al*. ROCr: visualizing classifier performance in R. *Bioinformatics* 2005;**21**:3940–1.
- R development core team. R: A Language and Environment for Statistical Computing, 2011.
- Aletaha D, Smolen J, Ward MM. Measuring function in rheumatoid arthritis: identifying reversible and irreversible components. *Arthritis Rheum* 2006;**54**:2784–92.
- Aletaha D, Ward MM, Machold KP, *et al*. Remission and active disease in rheumatoid arthritis: defining criteria for disease activity states. *Arthritis Rheum* 2005;**52**:2625–36.
- Scott D, Van Riel PC, van der Heijde DM, *et al*. Assessing disease activity in rheumatoid arthritis – The EULAR handbook of standard methods. On behalf of the EULAR Standing Committee for International Clinical Studies Including Therapeutic Trials – ESCISIT (Chairman: Smolen JS). Zürich: EULAR 1993.
- Smolen JS, Fleischmann R, Emery P, *et al*. The OPTIMA study of methotrexate and adalimumab: 78-week outcomes in early rheumatoid arthritis patients based on achieving a low DAS28 target after 26 weeks. *Ann Rheum Dis* 2011;**70**:259.
- Smolen JS, Emery P, Ferraccioli G, *et al*. Efficacy and safety of certolizumab pegol after incomplete response to DMARDs in RA patients with low moderate disease activity: results from CERTAIN, a phase IIB study. *Ann Rheum Dis* 2011;**70**:259.
- Felson DT, Smolen JS, Wells G, *et al*. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Arthritis Rheum* 2011;**63**:573–86.
- Felson DT, Smolen JS, Wells G, *et al*. American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann Rheum Dis* 2011;**70**:404–13.