

# A simplified disease activity index for rheumatoid arthritis for use in clinical practice

J. S. Smolen, F. C. Breedveld<sup>1</sup>, M. H. Schiff<sup>2</sup>, J. R. Kalden<sup>3</sup>,  
P. Emery<sup>4</sup>, G. Eberl, P. L. van Riel<sup>5</sup> and P. Tugwell<sup>6</sup>

**Objective.** The objective of this study was to verify the usefulness of a simple disease activity index (SDAI) for rheumatoid arthritis (RA).

**Methods.** The SDAI is the numerical sum of five outcome parameters: tender and swollen joint count (based on a 28-joint assessment), patient and physician global assessment of disease activity [visual analogue scale (VAS) 0–10 cm] and level of C-reactive protein (mg/dl, normal <1 mg/dl). Analysis initially focused on MN301, one of the three phase III clinical trials of leflunomide, in order to assess possible correlations between the SDAI and the Health Assessment Questionnaire (HAQ) and Disease Activity Score 28 (DAS 28). Results were then compared with the other two trials, MN302 and US301. A total of 1839 patients were evaluated. At baseline, 6 and 12 months, the SDAI, DAS 28, American College of Rheumatology (ACR) response criteria and mean HAQ scores were determined for each patient and compared by linear regression for significant correlation. The SDAI was compared qualitatively to the ACR 20% at 3, 6 and 12 months. The index was further validated by comparing the SDAI with survey results obtained from rheumatologists' evaluations of disease activity in test cases. The survey results included defining categorical changes in the SDAI indicating major, minor or no improvement in disease activity in response to treatment. Changes in total Sharp score at 6 and 12 months of treatment were determined for each of these categories of the SDAI and for comparable categories of the DAS 28.

**Results.** The mean SDAI calculated for patients at baseline in study MN301 was 50.06 (range 25.10–96.10) and was, respectively, 50.55 (range 22.10–98.10) and 43.20 (range 12.90–78.20) in studies MN302 and US301. In all three trials, the SDAI was correlated with a high level of statistical significance to the DAS 28 and HAQ scores at baseline, endpoint and change at endpoint. Patients achieving the ACR 20, 50, 70 or 90% response showed proportionate changes in the SDAI. Analysis of surveyed physician responses showed a significant association between the perception of disease activity and the SDAI, as well as changes in the SDAI. Qualitative analysis of radiographic progression at 6 and 12 months for patients showing either major, minor or no improvement of the SDAI showed correspondingly larger increases of the total Sharp score at 12 months.

**Conclusion.** The SDAI is a valid and sensitive assessment of disease activity and treatment response that is comparable with the DAS 28 and ACR response criteria; it is easy to calculate and therefore a viable tool for day-to-day clinical assessment of RA treatment. Overall results indicate that the SDAI has content, criterion and construct validity.

KEY WORDS: Simple disease activity index, Rheumatoid arthritis, Clinical practice.

University of Vienna and Lainz Hospital, Vienna, Austria, <sup>1</sup>University of Leiden, Leiden, The Netherlands, <sup>2</sup>University of Colorado, Denver, CO, USA, <sup>3</sup>University of Erlangen, Erlangen, Germany, <sup>4</sup>University of Leeds, Leeds, UK, <sup>5</sup>University Medical Centre, Nijmegen, The Netherlands and <sup>6</sup>University of Ottawa, Ottawa, ON, Canada

Submitted 27 December 2001; revised version accepted 2 July 2002.

Correspondence to: J. S. Smolen, Department of Rheumatology, Internal Medicine III, Vienna General Hospital, University of Vienna, Waehringerguertel 18–20, A-1090 Vienna, Austria. E-mail: josef.smolen@wienkav.at

Current therapy for the treatment of active rheumatoid arthritis (RA) favours aggressive treatment with disease-modifying antirheumatic drugs (DMARDs) with the goal of preventing or slowing permanent structural damage to the joints and limiting long-term disability. Inherent in this strategy is accurate monitoring of disease activity in order to follow disease progression and assess the effect of therapy [1, 2]. A variety of factors confound the process of measuring disease activity including the unpredictable course of RA and the varied clinical presentation of the disease in different patients [2–4]. Because of these factors, monitoring disease activity requires a composite evaluation of a variety of clinical parameters [1–7].

The selection of an index composed of outcome assessments for monitoring disease activity is governed by parameters sensitive to change, their predictive quality for disease status long-term, and whether the compilation of variables is comprehensive and the variables are not redundant [1, 8]. Both the European League Against Rheumatism (EULAR) and the American College of Rheumatology (ACR) as well as the World Health Organization/International League of Associations for Rheumatology (WHO/ILAR) have defined core sets of disease activity measures for RA with the goal of providing uniformity in the assessment of outcome in clinical trials [1, 5, 9]. The variables contained in both these sets include: tender and swollen joint counts, patient and physician global assessments of disease activity, acute-phase reactants, pain and assessments of physical disability. For RA trials of 1 yr or longer the inclusion of radiographic assessment in the evaluation of treatment outcome at study endpoint was specified.

The current standards of disease activity indices for clinical trials are the EULAR Disease Activity Score (DAS) [10], as well as the modified DAS 28 based on a 28-joint assessment [11], and the ACR 20% response criteria [12]. The EULAR response criteria [13] composed of categorical changes in the DAS or DAS 28 and the ACR 20% response indices are validated criteria for the assessment of treatment response [5, 11, 14, 15]. Comparative analysis of these criteria indicates a high level of agreement between the two methods in terms of their association with radiographic progression, the discriminating potential of each measure and the agreement with physician and patient assessments [14]. Each has proven viability and reliability in clinical trials. Nevertheless, limitations to each may restrict use in the daily practice of rheumatology and in the evaluation of treatment effect in RA.

The ACR 20% and the EULAR response criteria emphasize change in disease state and, therefore, are tools to assess clinically relevant improvement in disease activity. However, the nature of the calculation of the ACR 20% response criteria does not allow the measurement of actual disease activity [14] and does not enable comparison of one patient's absolute response with that of another one, which it is not designed for. Although designed to facilitate relative comparisons between treatment groups, such comparison is partly hampered

since the ACR response does not allow discernment of whether one group of patients had more active disease than another one at baseline. The EULAR criteria are based on the DAS score, which measures an absolute disease activity; however, the equation used to calculate the DAS score is relatively complex for immediate determination in the daily clinical setting and requires the use of a calculator.

It would be helpful, therefore, to obtain a simple disease activity index (SDAI) that retains the sensitive assessment characteristics of the DAS and ACR response criteria, yet is an intuitive, easy way to assess disease activity in daily clinical practice. It is based on a recently developed simple disease activity index for reactive arthritis (DAREA) [16], which consists of an arithmetical sum of five core set variables selected to be most suitable for such an index and proved to be valid, reliable and sensitive to change among patients with reactive arthritis. The current investigation tests a slightly modified version of the DAREA for RA. Validation of the SDAI was accomplished by the analysis of the leflunomide database of patients with active RA as this represents the largest compilation of RA patient data derived from controlled clinical trials evaluating the efficacy and safety of new DMARDs. In this study we assessed the value of this simply obtainable index by relating its changes to those of other established composite measures for RA, the ACR and DAS response, as well as to changes of the Health Assessment Questionnaire (HAQ) and of radiographic progression. The sensitivity to change of the variables selected for the SDAI was not re-evaluated here, since all these items are component measures within the core set of variables consistently shown to be clinically important, sensitive to change and reliable in investigations of study groups of the ACR, EULAR and OMERACT/ILAR/WHO [3–15]. Data obtained from one trial were validated using those from two other phase III trials.

## Methods

### *SDAI*

Calculation of the SDAI is the simple linear sum of the outcome parameters: tender joint count (TJC) and swollen joint count (SJC) based on a 28-joint assessment, patient global assessment of disease activity [PGA visual analogue scale (VAS) 0–10 cm], physician global assessment of disease activity (MDGA VAS 0–10 cm) and C-reactive protein (CRP in mg/dl) as shown below. The sensitivity to change of the individual variables selected for the SDAI was not re-evaluated here, since all these items are component measures within the core set of variables consistently shown to be clinically important, sensitive to change and reliable in investigations of study groups of the ACR, EULAR and OMERACT/ILAR/WHO [3–15]. The variables selected have also been shown to be sensitive to change in reactive arthritis [16]. An analogous index consisting of the arithmetic sum of such variables, DAREA, has been validated to reflect clinically important changes and be reliable in a clinical trial of reactive arthritis.

CRP rather than erythrocyte sedimentation rate (ESR) was employed since it is at least equally reliable in clinical trials [17] and, expressed as mg/dl, does not outweigh a laboratory variable in the score as would the ESR expressed as mm/h. Thus, the formula of the SDAI is as follows:

$$\text{SDAI} = \text{TJC} + \text{SJC} + \text{PGA} + \text{MDGA} + \text{CRP}$$

The range of SDAI values for this study is included in the results section.

### *Leflunomide database*

The SDAI was compared with the HAQ, DAS 28 and ACR response criteria that were determined for patients enrolled in one of the three phase III clinical trials, a multinational clinical trial of leflunomide (MN301), and validated by comparison with the HAQ, DAS and ACR response in the other two phase III clinical trials in the leflunomide database (MN302 and US301). The analysis presented in this study focuses on the entire patient populations treated during the individual phase III clinical trials: leflunomide *vs* placebo, *vs* sulphasalazine (MN301—6 months of treatment) [17]; leflunomide *vs* placebo, *vs* methotrexate (US301—12 months of treatment) [18]; and the comparative study of leflunomide *vs* methotrexate (MN302—12 months of treatment) [19]. Thus, data of all patients of individual trials, regardless of their treatment, were combined for this investigation. (In two of the trials analysed here, MN301 and 302, MDGA was originally assessed on a 5-point Likert scale, which was transposed to correspond to a 10-point VAS.)

The endpoint parameters determined in each trial included the ACR response criteria, composed of the clinical variables defining the criteria as indicated above, as well as rheumatoid factor (RF) and radiographic assessment of disease progression via either the Sharp or Larsen methods. The leflunomide database collected the core set of clinical outcome assessments, radiographic assessment of disease activity and an extensive library of demographic data for each treated patient. As such, the leflunomide database represents the largest collection of data concerning randomized controlled trials of DMARD therapy of RA patients with over 2241 cases randomized and 1339 patients treated with leflunomide. Inclusion in the leflunomide trials required a diagnosis of active RA based on the revised classification criteria of the ACR [20] and was of functional class I, II or III. Only patient data from phase III clinical trials were analysed in this investigation.

### *Comparison with validated disease assessments*

The SDAI was compared with the HAQ as well as the DAS 28 (shown below) [11] and the ACR response criteria [12], both of which include many of the same components.

$$\text{DAS 28} = 0.56\sqrt{\text{TJC}} + 0.28\sqrt{\text{SJC}} + 0.70[\ln(\text{ESR})] + 0.014(\text{general health})$$

The ACR 20% criteria [12] are defined as  $\geq 20\%$  improvement in TJC and SJC and  $\geq 20\%$  improvement in at least three of the following five parameters: patient self-assessed function (HAQ), PGA, MDGA, patient pain assessment and acute-phase reactant value measured by ESR or CRP.

### *Comparative analyses of the SDAI*

Disease activity was calculated using the SDAI, DAS 28 and ACR response criteria (20, 50, 70 and 90% response) for each

of the treated patients in the three studies (MN301, MN302 and US301) at baseline and study endpoint, and the change in activity at endpoint [intent-to-treat (ITT), last observation carried forward (LOCF) analysis]. The SDAI was also compared with the HAQ and modified HAQ (MHAQ, assessed in US301) scores in all treated patients. The SDAI, DAS 28 and HAQ scores constitute continuous variables and are amenable to simple regression analysis of linear association. Qualitative comparisons were made between the SDAI and ACR response criteria.

### *Comparison of survey results with the SDAI*

The overall rank order of disease activity determined by surveyed physicians (see below), as well as their mean assessments for disease activity classified as mild, moderate or severe, were compared with the SDAI ranking via simple linear regression to determine significant association between the assessments. The physicians' mean assessments of changes in the components of the SDAI, categorized according to no, minimal or major improvement, were also tested for significant association with sequentially ordered changes of the SDAI by linear regression analysis.

As a further test of the SDAI validity, the change in total Sharp score at 6 and 12 months' treatment for all patients in the three trials was determined for those who met either major, minor or no improvement categorical SDAI values determined by the results of the physician survey results outlined above. Similar analysis was conducted for categorical changes of the DAS 28 indicating either good (change in DAS 28 value  $> 1.2$ ), moderate (value  $> 0.6$  and  $\leq 1.2$ ) or no (value  $\leq 0.6$ ) improvement. Construct or convergent validity are evaluated by comparison with the accepted standards of assessing disease activity and treatment outcome (DAS 28, ACR response and HAQ) and physician surveys, respectively.

### *Rheumatologist survey*

To test further the usefulness of the SDAI as a measure of disease activity, an informal survey among 21 European and North American rheumatologists was conducted. These rheumatologists were selected on the basis of their clinical expertise—junior and senior rheumatologists (about 50% of each) were included among them; a number of them are actively involved in rheumatology training programmes. Survey participants were provided with RA cases containing the data for the five SDAI components blinded to the SDAI derivation and its value and were asked to: (i) rank the cases by disease activity (20 cases evaluated); (ii) categorize the disease activity for 20 cases in terms of mild, moderate or severe disease (scored 1, 2 or 3, respectively); and (iii) categorize a change in disease activity as no, minimal or major improvement (25 cases evaluated and scored on an ordinal scale of 1, 2 or 3, respectively). All assessments returned were included in the calculations. Mean assessments were calculated by summing the numbers obtained when a particular grade was multiplied with the number of physicians assigning that specific grade and dividing by the total number of physicians responding.

### *Statistical analysis*

Continuous variables (SDAI, DAS 28 and HAQ) were compared via linear regression with the SDAI assigned as the independent variable. The probability of a significant association between the independent and dependent variables was defined as  $P < 0.05$ .

## Results

### Patient demographics and the SDAI at baseline

Patient demographics have been reported previously for each of the three phase III clinical trials [17–19]. The total sample size analysed in this investigation comprised 1839 patients with RA. Mean patient ages observed in the respective trials ranged between 53 and 59 yr with the majority <65 yr. In addition, the patients were predominantly women with a mean disease duration of 4–8 yr (~40% with disease duration of ≤2 yr). All patients had active RA and the large majority was

classified as functional class II or III [21]. Table 1 summarizes patient demographics and the baseline values of selected clinical parameters including mean value of the SDAI and DAS 28. The mean SDAI calculated for all patients at baseline in study MN301 was 50.06 (range 25.10–96.10) and, respectively, 50.55 (range 22.10–98.10) and 43.20 (range 12.90–78.20) in studies MN302 and US301. The ranges of TJC, SJC, PGA, MDGA and CRP were 0–28, 0–28, 0–10 cm, 0–10 cm and 0.1–23.7 mg/dl, respectively, for the intent-to-treat population in all treatment groups for all trials.

TABLE 1. Leflunomide phase III trials: patient demographics and baseline clinical data

	LEF	PL	SSZ	MTX	Totals
<b>MN301</b>					
<i>n</i>	133	92	133	0	358
Mean age (yr)	58	59	59	...	...
% <65 yr	68	71	62	...	...
% Women	76	75	69	...	...
Mean disease duration (yr)	8	6	7	...	...
Disease duration ≤2 yr (%)	38	45	42	...	...
ACR functional class II, III	92	97	95	...	...
Mean tender joint count*	18.8	16.3	16.7	...	...
Mean swollen joint count*	16.2	15.8	15.3	...	...
Physician assessment	3.6	3.5	3.5	...	...
Patient assessment	3.7	3.6	3.6	...	...
CRP (mg/dl)	4.5	4.1	3.4	...	...
SDAI (all patients)				Mean 50 ± 14	Range 25–96
DAS 28 (all patients)				Mean 7.0 ± 0.833	Range 5–9
<b>MN302</b>					
<i>n</i>	501	0	0	498	999
Mean age (yr)	58	...	...	58	...
% <65 yr	70	...	...	70	...
% Women	71	...	...	71	...
Mean disease duration (yr)	4	...	...	4	...
Disease duration ≤2 yr (%)	44	...	...	43	...
ACR functional class II, III	93	...	...	93	...
Mean tender joint count <sup>a</sup>	17.2	...	...	17.7	...
Mean swollen joint count <sup>a</sup>	15.8	...	...	16.5	...
Physician assessment	3.5	...	...	3.6	...
Patient assessment	3.6	...	...	3.6	...
CRP (mg/dl)	4.3	...	...	4.0	...
SDAI (all patients)				Mean 51 ± 14	Range 22–98
DAS 28 (all patients)				Mean 7.0 ± 0.790	Range 5–9
<b>US301</b>					
<i>n</i>	182	118	0	182	482
Mean age (yr)	54	55	...	53	...
% <65 yr	78	82	...	81	...
% Women	73	70	...	75	...
Mean disease duration (yr)	7	7	...	7	...
Disease duration ≤2 yr (%)	39	33	...	40	...
ACR functional class II, III	86	91	...	87	...
Mean tender joint count <sup>a</sup>	15.5	16.5	...	15.8	...
Mean swollen joint count <sup>a</sup>	13.7	14.8	...	13.0	...
Physician assessment	6.1	6.2	...	5.9	...
Patient assessment	5.6	5.8	...	5.4	...
CRP (mg/dl)	2.1	2.5	...	1.88	...
SDAI (all patients)				Mean 43 ± 13	Range 12–78
DAS 28 (all patients)				Mean 6.3 ± 1.072	Range 3–9

<sup>a</sup>Based on a 28-joint count.

LEF, leflunomide; PL, placebo; SSZ, sulphasalazine; MTX, methotrexate.

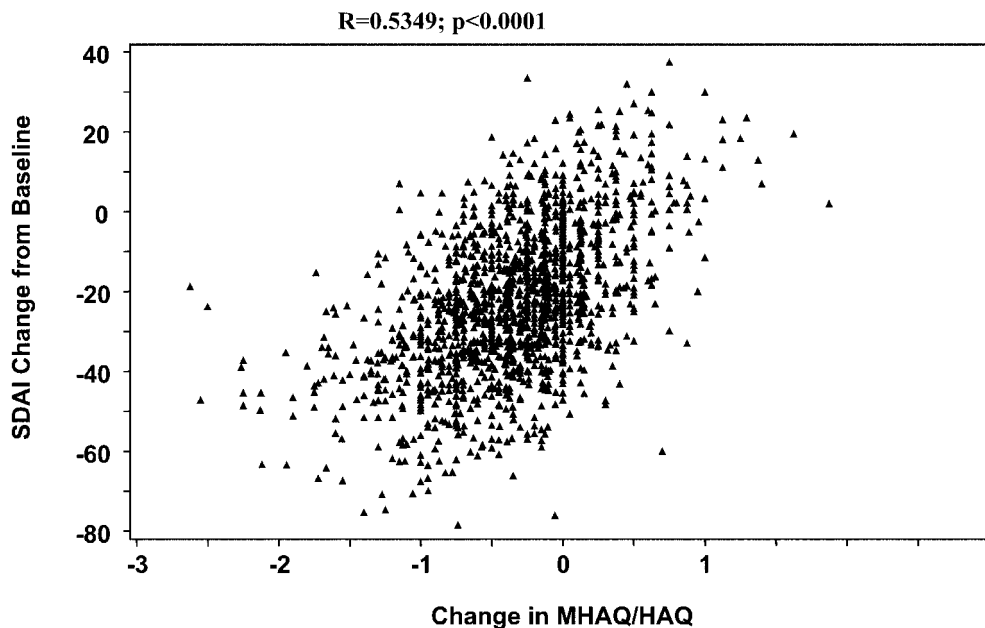


FIG. 1. Scatter plot of pooled patient data (see Table 1) showing the change in SDAI compared with the change in HAQ [and MHAQ (US301)] at 12 months of treatment (MN303, MN302 and US301). Change in the SDAI was significantly related to the change in HAQ in response to treatment ( $r=0.5349$ ;  $P < 0.0001$ ).

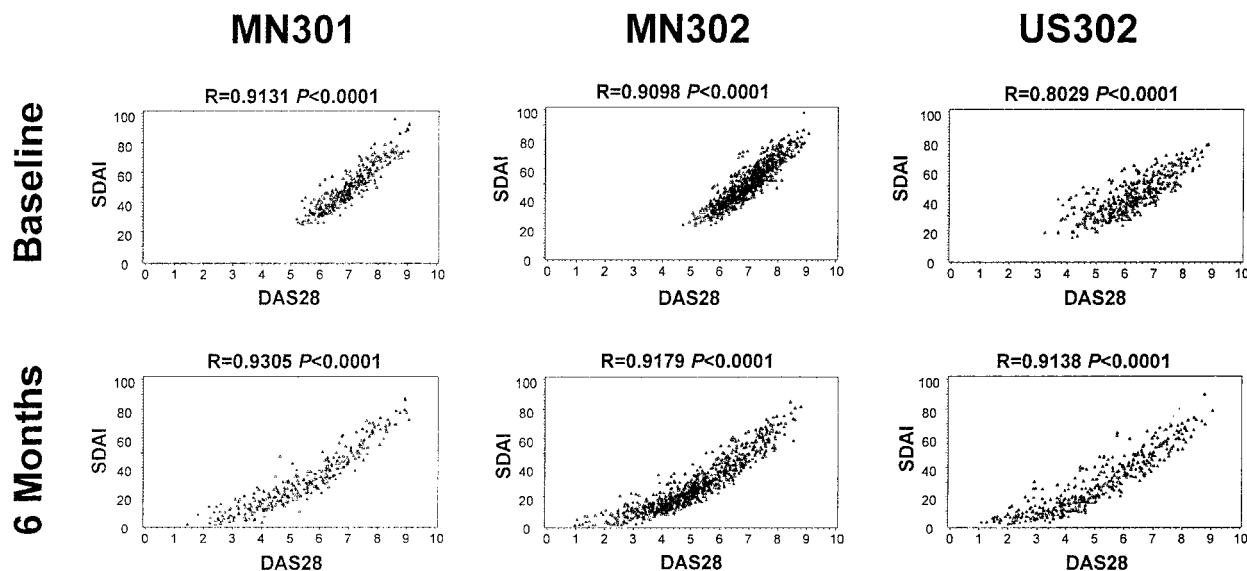


FIG. 2. Scatter plots of the DAS 28 vs the SDAI at baseline and 6 months of treatment. All linear regression analyses indicate a highly significant degree of correlation between the SDAI and DAS 28. The numbers of patients for each trial are shown in Table 1.

#### *Correlation and validation of the SDAI with the HAQ*

Comparison of the SDAI with the HAQ in study MN301 using linear regression revealed a significant linear relationship of change in the SDAI and change in the HAQ between baseline and 6 months ( $r=0.56$ ,  $P < 0.0001$ ). This correlation was validated in studies US301 and MN302 comparing changes in the SDAI with changes in HAQ or transformed MHAQ between baseline and month 12 ( $r=0.57$  and  $r=0.48$ , respectively,

$P < 0.0001$ ). At all time points, at baseline, and at 6 and at 12 months, respectively, there was also a linear relationship between both parameters showing a significant correlation (MN301:  $r=0.46$  at baseline and  $r=0.63$  at 6 months; US301:  $r=0.44$  at baseline and  $r=0.66$  at 12 months; MN302:  $r=0.36$  at baseline and  $r=0.53$  at 12 months;  $P < 0.0001$  for all analyses).

When physician global assessment was replaced by patient pain assessment (as originally done in the DAREA), the correlations between change of this

revised SDAI and change in the HAQ were almost identical (MN301:  $r=0.57$ ; US301:  $r=0.56$ ; MN302:  $r=0.49$ ,  $P < 0.0001$  for all analyses). When the SDAI was further modified to exclude CRP, the changes of such modified SDAI were again similarly significantly correlated to changes of the HAQ/MHAQ (MN301:  $r=0.56$ ; US301:  $r=0.56$ ; MN302:  $r=0.47$ ;  $P < 0.0001$  for all analyses).

Thus, there was a linear relationship between the SDAI and HAQ/MHAQ as well as between changes in the SDAI and HAQ/MHAQ in all three studies at all points in time confirming the validity and usefulness of the SDAI. Moreover, exchange of the physician's global assessment of disease activity as a component of the SDAI by patient's pain assessment, the component of the DAREA replaced in the SDAI by physician's global assessment, did not change the correlations.

For simplicity, the change in HAQ/MHAQ at 12 months in the compiled studies compared with the change in SDAI is shown in Fig. 1 and was significantly ( $r=0.5349$ ;  $P < 0.0001$ ) related, a reflection of disease improvement or worsening in conjunction with DMARD or placebo treatment.

When the SDAI was modified to exclude CRP, the changes of such modified SDAI were again similarly significantly correlated to the changes of HAQ/MHAQ (MN301:  $r=0.56$ ; US301:  $r=0.56$ ; MN302:  $r=0.47$ ;  $P < 0.0001$  for all analyses).

#### Correlation and validation of the SDAI with DAS 28

Figure 2 shows scatter plots of the paired SDAI values and corresponding DAS 28 scores at baseline and 6 months of treatment in each of the three trials. Linear regression analysis indicated a linear correlation between the SDAI and DAS 28 at baseline with high  $r$  values for MN301 ( $r=0.9131$ ), MN302 ( $r=0.9098$ ) and US301 ( $r=0.8029$ ) that proved highly significant with a  $P < 0.0001$ . Similarly, highly significant ( $P < 0.0001$ ) linear association was determined for the correlation between the SDAI and DAS 28 at 6 months: MN301 ( $r=0.9305$ ), MN302 ( $r=0.9179$ ) and US301 ( $r=0.9138$ ). The changes in SDAI compared with the changes in DAS 28 are shown for all three studies in Fig. 3 and confirm the highly significant correlations between the scores.

#### Comparison of the SDAI with the ACR response criteria

The ACR response criteria are not defined as a continuous variable and therefore require non-parametric methods of statistical analysis, which were not conducted for this investigation. However, Fig. 4 shows a comparison of the ACR 20, 50, 70 and 90% response with the corresponding change in SDAI at 3, 6 and 12 months for the pooled clinical data. In qualitative terms, the graph shows a proportionately greater change in the SDAI for the increasingly stringent ACR 20 to 90% response criteria, although it should be noted that

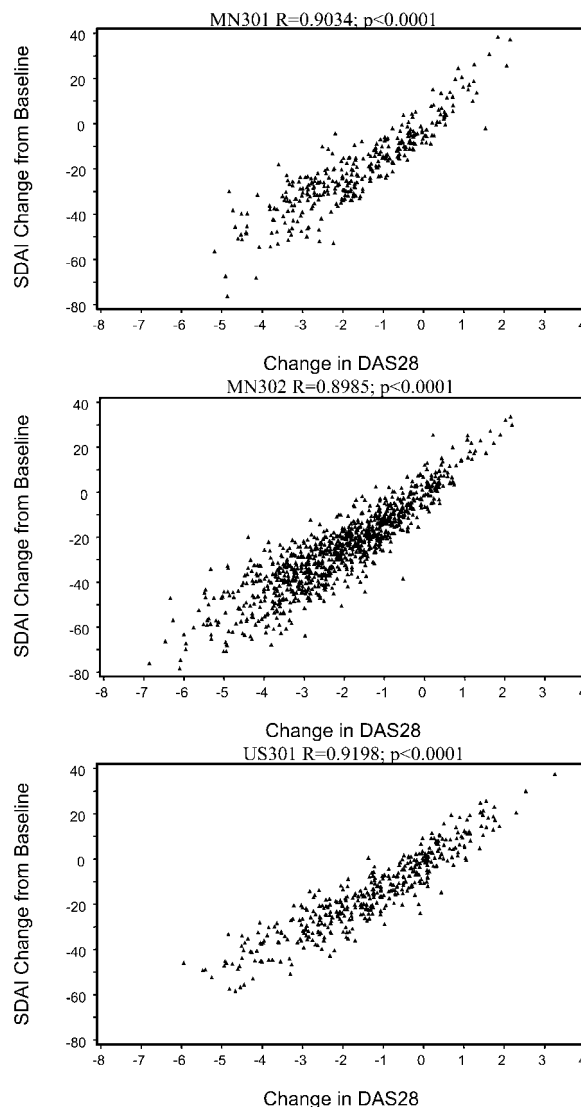


FIG. 3. Scatter plots of the change in DAS 28 vs the change in SDAI at 6 months treatment. All linear regression analyses indicate a highly significant degree of correlation between the SDAI and DAS 28. The numbers of patients for each trial are shown in Table 1.

patients achieving  $\geq 20\%$  response include  $\geq 50$ ,  $\geq 70$  and  $\geq 90\%$  responders as well.

#### Comparison of the SDAI, DAS, HAQ and ACR response

An improvement of 0.6–1.2 in the DAS score, 0.22 in the HAQ and  $\geq 20\%$  improvement of the ACR response criteria has been determined to be clinically meaningful in terms of assessing disease activity [12, 13, 22]. Comparison of the change in SDAI and DAS 28, HAQ, and ACR 20% response criteria at 3, 6 and 12 months of treatment (Fig. 5) indicates a fairly consistent and proportional change in the SDAI at the three time points. Changes in the SDAI were consistent ( $-16 \pm 5$  or 6) at all time points compared with a change in DAS 28

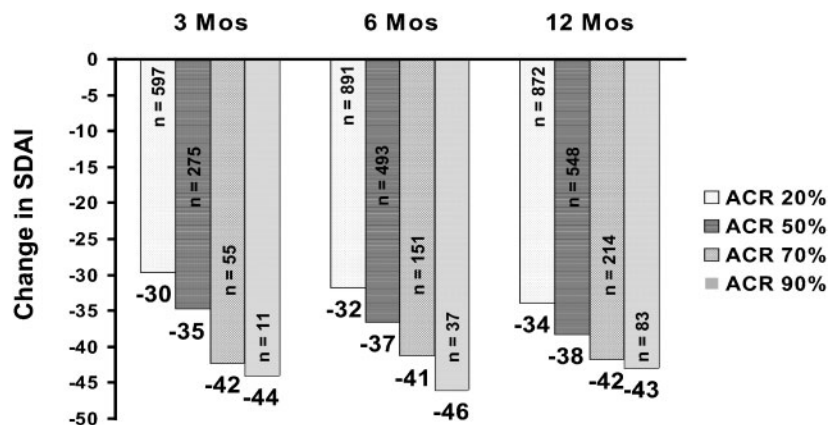


FIG. 4. Pooled ACR responders for phase III clinical trials of leflunomide over 3, 6 and 12 months of treatment. ACR response at increasingly greater levels shows proportionate changes in the SDAI.

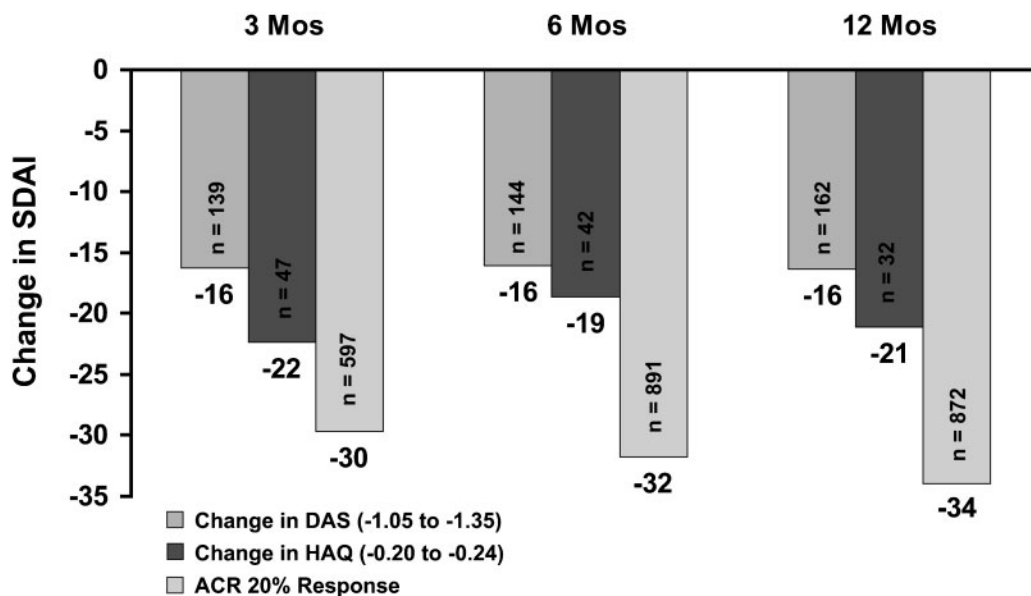


FIG. 5. Change in the SDAI compared with the change in DAS 28 ( $-0.5$  to  $-0.7$  and  $-1.05$  to  $-1.35$ ), change in HAQ ( $-0.20$  to  $-0.24$ ) and ACR 20% response (pooled data). Change in the SDAI at each time point is consistent in each category.

of between  $-1.05$  and  $-1.35$ , arbitrarily defining a small interval around the clinically meaningful improvement of 1.2 for the pooled clinical data (Fig. 5). Similarly, changes in DAS 28 of  $-0.6$  ( $-0.5$  to  $-0.7$ ) related to an SDAI change of  $-9$ .

There were no significant variations between changes in the SDAI and changes in HAQ/MHAQ at 3, 6 and 12 months ( $-22 \pm 12$ ,  $-19 \pm 12$ , and  $-21 \pm 17$ , respectively). The interval of HAQ change shown ( $-0.20$  to  $-0.24$ ) is relatively narrow and defines a clinically significant reduction in the HAQ [13, 22]; considering that the majority of patients treated with active compounds showed greater than  $-0.22$  improvement in HAQ scores [23], the sample size is limited. Changes in the SDAI slightly increase with respect to time ( $-30 \pm 11$  at 3 months,  $-32 \pm 12$  at 6 months, and  $-34 \pm 13$  at

12 months) and ACR 20% response and reflect the increasing number of treatment responders in higher categories of the ACR response criteria. (It should be noted that the combination of HAQ and MHAQ data may not be fully appropriate, but the data obtained are similarly valid for HAQ and MHAQ separately.)

#### Survey results

Figure 6 presents a direct categorization of 20 RA cases of increasing SDAI values ranked by the surveyed physicians [ $n=21$  (mean assessment)]. Statistical analysis verifies a highly significant ( $P < 0.0001$ ) linear correlation ( $r=0.8908$ ) between the two sets.

The number of physicians categorizing the same 20 RA test cases as mild, moderate or severe is shown in Fig. 7A. These cases were selected according to the

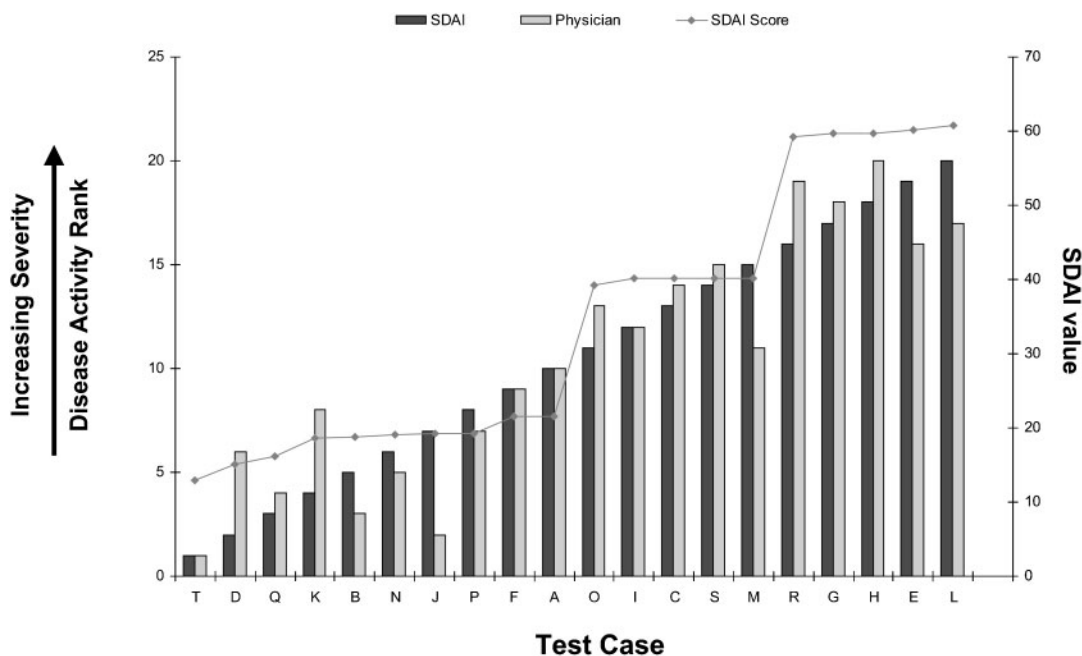


FIG. 6. Histogram of test case (A–T) disease activity, ranked according to increasing SDAI score, and physician assessment (left-hand axis). The dark bars indicate the rank of disease activity according to the SDAI, the light bars indicate the rank of disease activity according to the participating physicians. The diamonds, connected by a line, indicate each individual patient's actual SDAI score (right-hand axis). Note that the 10 patients with the lowest SDAI scores were also the patients with the 10 lowest ranks according to the physicians' assessments and the five patients with the highest scores had the highest ranks for disease activity. Physicians were blinded to the score and the method employed and only had information on the values of the individual components of the SDAI. Line graph shows the actual SDAI score (right-hand axis).

SDAI (<20, 40, 60), which were unknown to the assessors. None of the physicians scored any patient with an SDAI <20 as severe. The majority (70%) of physicians scored patients with an SDAI of ~40 (patients O, I, C, S, M) as having moderate disease activity (neither mild nor severe), and patients with an SDAI of 60 were assessed as having severe disease activity. In Fig. 7B we show a comparison of the actual numerical values of the SDAI for each patient (shown as a line linking all patients in Fig. 7B) with the means of the assessments by the physicians' assessments; the data reveal close parallelism of the two evaluations (Fig. 7B). Statistical comparison between the mean assessment and the SDAI proved a highly significant linear ( $r=0.9434$ ) association between the two parameters ( $P < 0.0001$ ). Moreover, these analyses compared with the physicians' categorization of the disease as mildly, moderately or severely active. The cases presented in this analysis indicate that an SDAI <20 is considered mild, between 20 and 40 moderate, and > 40 severe activity of RA.

The final component evaluated by the survey was the ranking of changes in the SDAI. Figure 8 presents both the change from baseline of the SDAI and the physician mean assessment of improvement. The surveyed rheumatologists were asked to rate different changes of clinical activity as major improvement (rated 3), minor improvement (rated 2) or no improvement (rated 1). The change in SDAI for the individual patient was

compared with the mean result of the physician assessment of improvement. There was a highly significant inverse linear correlation ( $P < 0.0001$ ) between the mean assessed improvement and the change in SDAI indicating that large changes in the SDAI were judged as showing greater improvement in disease activity compared with smaller changes. Based on the mean assessment of disease activity, a reduction in the SDAI of between -10 and -21 was judged as minor improvement, while a change of -22 or more was considered major improvement by a majority of the physicians.

#### *Clinical improvement defined by the SDAI*

All data taken together—i.e. changes in the SDAI compared with DAS, HAQ and physician survey results regarding disease activity and change in response to treatment—suggest that: improvement in the SDAI of -22 or more represents major improvement, while improvement of between -10 and -21 represents minor improvement. If one examines these results in relation to the DAS, a change of -0.6 in the DAS 28 corresponded to an improvement of -9 in the SDAI, a change of -1.2 in the DAS corresponded to a change of -16 in the SDAI, and a change of -0.22 in the HAQ was associated with a change of -19 in the SDAI (Fig. 5). Moreover, all data together reveal that an absolute SDAI value of 5–20 relates to mild disease activity, while an SDAI of 21 to 40 corresponds to moderate disease



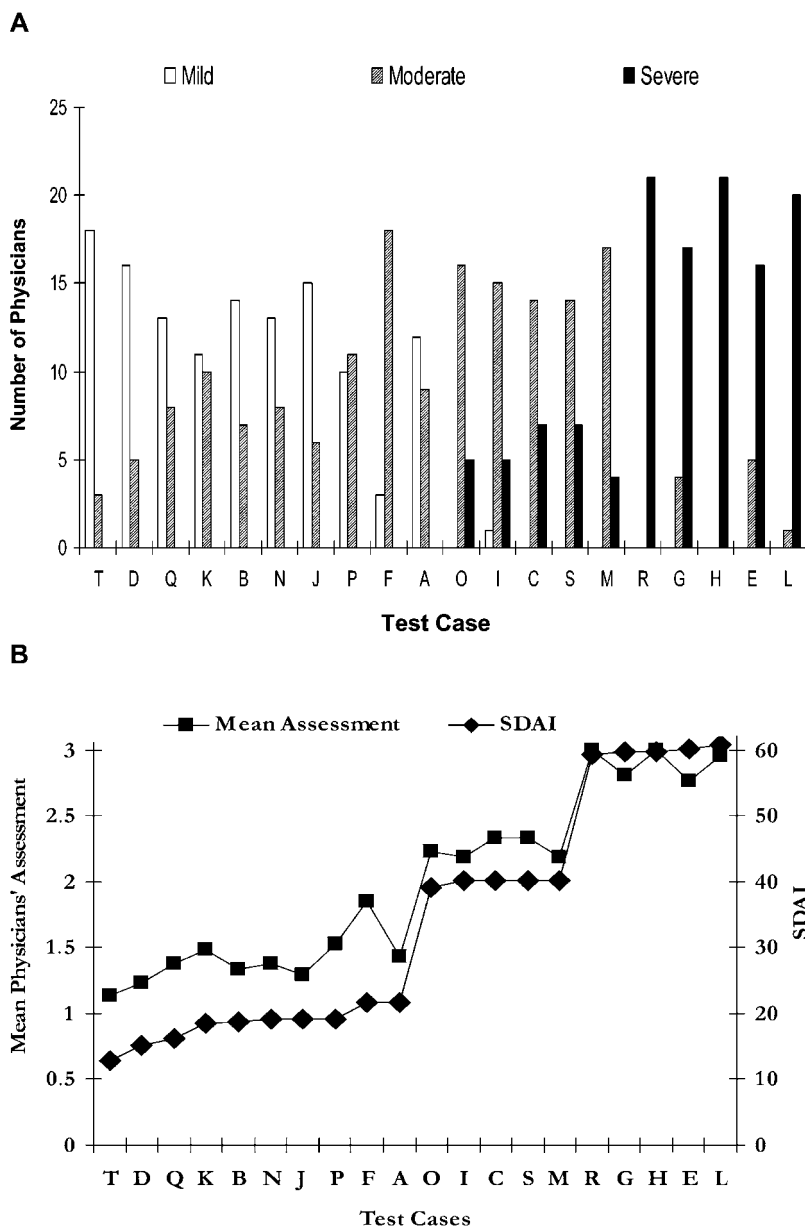


FIG. 7. Histogram of disease severity showing the number of physicians judging 20 test cases (7A) according to disease severity (mild, moderate or severe; cases A–T). The physicians’ mean assessment of severity is shown in 7B with the SDAI score for each test case (right-hand axis); the actual SDAI scores for each patient are indicated by the diamonds while the mean assessments (mild = 1, moderate = 2, severe = 3) were shown as squares, and both connected by a line for better comparability of patients. The patients assessed here for mild, moderate and severe disease activity were the same as those ranked for increasing disease activity in Fig. 6.

activity, and finally, an SDAI of > 40 is associated with severe disease activity.

*Correlation of changes in the SDAI with radiographic changes*

The change in total Sharp score for patients fulfilling the categorical changes of major, minor or no improvement in the SDAI or good, moderate or no response to treatment in the DAS 28 is shown in Fig. 9. Major improvement in the SDAI (decrease by at least 22) at

12 months of treatment corresponded to a mean increase of total Sharp score of 1.1. Moderate improvement (decrease in SDAI between 10 and 21) corresponded to a mean increase of 1.9 in total Sharp score, and for patients showing no improvement in the SDAI (decrease of 9 or less) the total Sharp score increased the most with an observed progression of 3.2 points, also at the 12-month time point. Similar to patients with no improvement according to the SDAI, non-responders according to the DAS had an increase in Sharp score of 3.2, while good DAS responders had a much lower

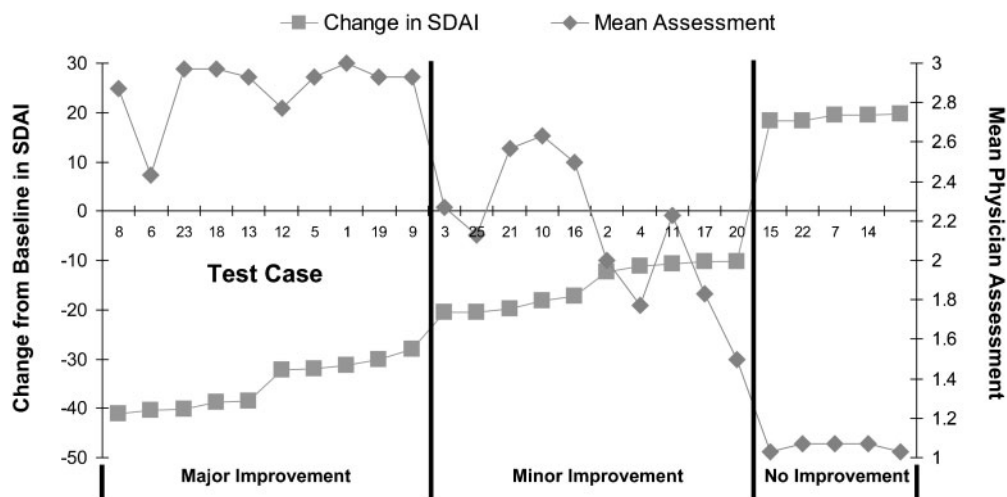


FIG. 8. SDAI change from baseline for 25 test cases (cases 1–25) with the physician global assessment (major, minor or no improvement) of each case. These cases were different from those shown in Figs 6 and 7, since the data provided here related to changes of the individual variables used to derive the SDAI with the doctors blinded for the score and the calculation. The data shown constitute mean values of assessing physicians (diamonds) who indicated no improvement (score of 1), minor improvement (score of 2) or major improvement (score of 3). Note that patients with the highest improvement in SDAI (left-hand part of the graph) also were regarded as those with major improvement, whereas most of those with changes of SDAI between -10 and -22 were regarded to have minor improvement.

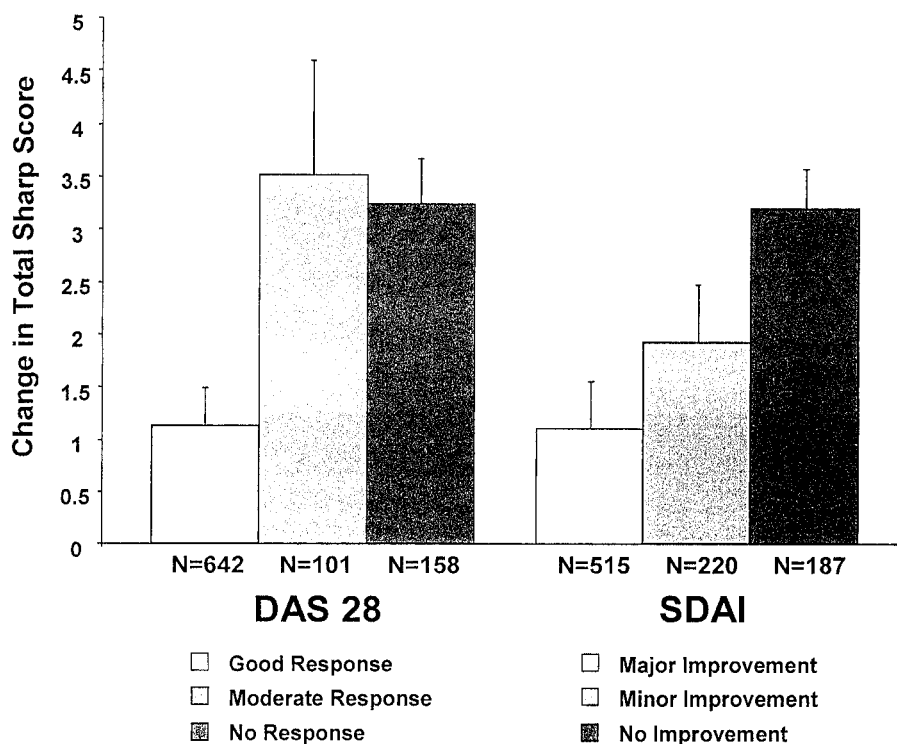


FIG. 9. Change in total Sharp score compared with categorical changes in the SDAI and DAS 28. A generalized increase in the total Sharp score was observed for good, moderate and non-responders employing the DAS 28. Similarly, changes in the SDAI indicating major, minor and no improvement of disease activity corresponded to comparable increases in total Sharp score. Data presented in graph from pooled patient data.

progression (1.1), which was similar to that of patients with a major response by SDAI. When the Larsen score was employed, there were smaller changes among

patients with major SDAI improvement than among those with no improvement confirming the results obtained using the Sharp score (data not shown).

## Discussion

The SDAI formulated and validated here combines swollen joint counts, tender joint counts, patient global assessment, physician global assessment and CRP (mg/dl) in a simple numerical summation. The results reveal a highly significant association between the SDAI and the DAS 28, as well as the HAQ scores of patients with active RA treated in several phase III clinical trials of leflunomide. The SDAI was primarily validated in the second and third of these trials by confirming its significant correlation with functional assessment by HAQ and with the DAS 28 and particularly by showing a linear relationship between the change in SDAI and the change in HAQ and DAS 28 in the first of the trials. In addition, the qualitative comparison of the SDAI with ACR response shows that proportionately greater changes in the SDAI correspond with the ACR response criteria at 20, 50, 70 and 90% levels. Further validation is provided by evidence derived from surveys of rheumatologists that correlate the clinical perception of disease activity with the SDAI and disease improvement with changes in SDAI. Finally, the SDAI proved useful in a cohort of patients followed in a rheumatic disease clinic (data not shown). Thus, based on a recently validated index of disease activity for reactive arthritis [16], the SDAI evaluated in this study exhibits potential application as a new and easily applicable index for the assessment of disease activity in the day-to-day care of RA.

At baseline, all patients analysed had active disease, as can be seen by their high DAS scores (means of 6.3–7.0). Patients in each of the trials had mean SDAI values of 43–51. Results of the physician survey indicated that an SDAI value of  $> 40$  constitutes high disease activity, an SDAI of 20–40 indicates moderate RA activity, and an SDAI of  $< 20$  indicates mild disease. These values correspond well with the categorical levels of disease activity defined for the DAS 28.

A change in the SDAI of  $-22$  or more was observed to represent major improvement, while a change of  $-10$  to  $-22$  suggested moderate improvement. A change in the SDAI of  $-10$  is very close to the value of  $-9$  associated with a change of  $-0.6$  of the DAS 28 that is associated with moderate improvement in response to treatment [11]. The major improvement in response to treatment indicated by a change of more than  $-22$  in the SDAI value surpasses the SDAI value of  $-16$  associated with a change of  $-1.2$  in the DAS 28 score, indicating major clinical improvement [9]. However, major improvement in the SDAI closely matches the SDAI value ( $-20.7$ , average value observed at 3, 6 and 12 months treatment) observed for a clinically significant improvement of the HAQ score indicated by a  $-0.22$  change from baseline [22]. An even greater change in the SDAI was associated with the ACR 20% response; however, ACR 20% responders also include patients who have much higher ACR responses.

The content validity of the SDAI is supported by the observation that assessed parameters include physician and patient assessments and a laboratory variable; these

are representative of the core set of variables needed in outcome assessment of RA [1]. The criterion validity of the SDAI is confirmed by the high degree of statistically significant correlation with the validated disease activity score (DAS 28) in terms of the baseline, endpoint and change at endpoint values. In addition, there is a significant association between changes in the HAQ/MHAQ and changes in SDAI, though the  $r$  values do not imply a very strong linear association. Further confirmation is observed in qualitative comparisons of the SDAI and ACR response criteria, which indicate proportionate changes in the SDAI with increasing levels of ACR response criteria.

The construct validity is exemplified by the distinct parallels among physician surveys, indicating a strong correlation between increasing activity and increasing SDAI and between greater changes in the SDAI with greater degrees of clinical improvement from the clinician's perspective.

All components of the SDAI are variables of the EULAR and ACR core set of outcome assessments [13, 22]. Joint counts and global assessments are given comparable weight in the SDAI. Physician's global assessment had been converted from a 5-point Likert to a 10-point VAS scale in two trials. Although this may have led to a somewhat lesser sensitivity to change in these trials, the similarity of the results obtained among all trials suggests that this does not confound the conclusions drawn. To reduce the potential that joint counts might dominate the index, the 28-joint count has been employed [13, 24, 25]. Joint counts are correlated with the current state of disability and swollen joints are predictive for radiographically evident disease progression [1].

The general concept was simplicity whilst retaining methodological rigour. Patient and physician global assessments of disease activity are both correlated with future disability [26] and are sensitive measures of change in clinical trials [27]. Although they are subjective evaluations, both are included in the ACR and WHO/ILAR core set of disease activity assessments [1, 9], in part because of their face validity, but also for their sensitivity to change in response to treatment in clinical trials [28, 29]. Pain assessment was not included because it is also reflected in patient global assessment. Although the importance of the HAQ/disability score is widely acknowledged, clinicians have failed to incorporate it in daily practice. Therefore it was not included for the purpose of this study. A further reason not to include it was to be able to use it as an independent variable for the assessment of the utility of the SDAI. In fact, this was of particular importance given that the components of the SDAI are also contained in the DAS and the ACR response criteria and that the HAQ also constitutes part of the ACR response evaluation.

Whereas swollen joint counts and physician global assessments are physician-derived assessments, tender joint counts (sometimes interpreted as a physician assessment) and patient global assessments depend upon the patient's perception of disease activity. Thus

the SDAI gives weight both to patient and physician assessments of disease activity. Importantly, when a revised SDAI containing patient pain assessment in place of physician's global assessment, as originally used in the DAREA [16], was evaluated, the correlation with the HAQ was virtually identical. Inclusion of two joint counts was done for the same reasons as they were recommended for the core set [6, 9–12].

The remaining assessment in the SDAI is a measure of acute-phase response. Both CRP and ESR are non-specific markers of inflammation that correlate with RA disease activity [15], as well as being sensitive measures of change in response to treatment in clinical trials [30]. Current research indicates the predictive value of elevated CRP levels for future development of progressive disease [31–33]. CRP levels are also less influenced by drug toxicity, age and sex and have proven to be a dynamic and direct measure of the acute-phase response [16]. Moreover, they are also recommended in the core sets of the ACR, EULAR and OMERACT. Calculation of the SDAI using CRP employs the mg/dl instead of the g/l concentration to avoid overemphasizing the laboratory assessment in the overall score [16]. In general, however, CRP levels did not add much to the relationship of the SDAI with the HAQ. When a modified SDAI containing all other variables except for CRP was assessed, similar correlations between such modified SDAI and the HAQ were observed as for the unmodified SDAI; moreover, there was a significant, albeit weak relationship between such modified SDAI and CRP levels (data not shown). Mean CRP levels amounted to 3.56 mg/dl at baseline with a change to endpoint of 1.52 mg/dl, across all studies.

Weighting was discussed but weighting had not been shown to be of additional benefit in predicting future morbidity and mortality [4, 5, 12]. In the calculation of the SDAI, joint counts carry the biggest weight, at least in active disease. In fact, future disability has been shown to be related to joint count [34]. Thus, although the SDAI is mainly dominated by the joint counts in active disease, the decision was against weighting up or down of individual variables. The good correlation of SDAI with HAQ scores, but also with the ACR response (where joint counts predominate somewhat less) and the DAS score support this decision.

Patient self-assessments of physical function, specifically the HAQ, have proven to be very sensitive measures and predictors of long-term disability [35]. In addition, there is a distinct link between the HAQ assessment of disability and progression of long-term damage [36]. Indeed, results from clinical trials of leflunomide indicate that the HAQ is a sensitive measure of early change with significant improvement in response to treatment evident at 1 month [17]. Observations presented in this investigation confirm a significant association of the HAQ score with the SDAI at baseline and at 6 and 12 months of treatment, as well as the correlation of changes in SDAI with changes in HAQ/MHAQ, thus reaffirming the validity of the SDAI. However, the SDAI does not provide the same information as

the HAQ (which measures function rather than disease activity) and thus cannot substitute for the HAQ. In fact, the SDAI is not meant to substitute for the HAQ in daily practice, but rather ought to provide an easily assessable means to calculate disease activity.

A correlation between the SDAI and both the DAS 28 and ACR response criteria may not be surprising given that the SDAI uses a subset of the same component parameters. However, the high level of correlation between the SDAI and the DAS 28 was not necessarily expected, and the correlation with the HAQ and changes in HAQ following treatment confirm the validity of the SDAI. Importantly, the SDAI value observed to represent a major improvement or change in disease activity determined by the survey results presented in this study matched almost exactly the value determined to correspond to a clinically significant change in the HAQ [22]. Further support for the validity of the SDAI as an index of disease activity is the observed very small increase in radiographic progression of RA in patients showing major SDAI improvement, while a much higher increase in radiographic progression occurred in patients who had no response or improvement following treatment.

Measures of RA disease activity should also reflect the clinical perception of disease activity, meaning that a rheumatologist's identification of a severe case of RA should be matched by an equally high value of the index. Similarly, the index should also be sensitive to change in disease status indicating a response, or lack of response, to treatment that is clinically perceptible and proportionate to the change in index from a clinical perspective. The results of the clinical survey reported in this study clearly indicate a high level of correlation between clinical rheumatologists' perception of disease activity and the SDAI, as well as between changes in SDAI and improvement or worsening of disease. A comparison of absolute SDAI values and changes in SDAI with treatment against the DAS, HAQ, ACR response and radiographic progression, as well as physician survey results, indicates clear achievement of face/content, criterion and construct validity essential for the validation of the SDAI as a new index of disease activity.

Employing the leflunomide database to validate the SDAI highlights the significant impact such resources could have on the study of RA. The leflunomide database represents the largest compilation of RA patient data hitherto published from prospective randomized controlled trials of DMARDs including demographics, disease duration and activity (at baseline and following treatment), clinical features and laboratory parameters—all collected under consistent and comparable conditions during clinical trials. Subanalysis of such a database, with its large patient numbers, may allow the determination of unique and significant disease characteristics not easily discriminated by studies of smaller size. The utility of the large database is exemplified by the fact that the large database allowed the SDAI to be assessed in one study and validated by comparison with the other two clinical trials. Plans are in progress to test the SDAI with

simulated patient profiles in order to assess the value of the score in making therapeutic decisions. The next phase will involve testing it in real clinics longitudinally. We will also ask rheumatologists if the SDAI gives them more information than they currently have in the follow-up of their patients. Furthermore, we will assess if one can obtain similar results as well as acceptability from a score consisting of smaller numbers of variables.

In summary, the results of this investigation illustrate that it is possible to obtain a sensitive index by forming an arithmetic sum of core set variable data. The SDAI is a simple and effective measure of disease activity in RA. The SDAI is significantly correlated with the DAS 28 and the HAQ, both validated and sensitive measures of disease activity. In addition, the ACR response at increasingly stringent levels is mirrored by proportionate changes in the SDAI. Moreover, increasing improvement by the SDAI is associated with decreasing radiographic progression. Given these observations, the SDAI may be a viable supplement to the DAS 28 or ACR criteria. Although, given its validity, the SDAI could be used in clinical trials of new therapies, the ideal application for the SDAI is in clinical practice for the day-to-day assessment of RA treatment effect, validating current therapy or indicating the need to revise or adapt treatment for progressing disease.

## References

1. 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. *Arthritis Rheum* 1993;36:729–40.
2. van Gestel AM, van Riel PLCM. Evaluation of early rheumatoid arthritis disease activity and outcome. *Baillière's Clin Rheumatol* 1997;11:49–63.
3. Boers M, Tugwell P, Brooks PM. Progress towards optimal trial end-points in rheumatoid arthritis. *BioDrugs* 1997;7:40–50.
4. Goldsmith CH, Boers M, Bombardier C, Tugwell P. Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. *J Rheumatol* 1993;20:561–5.
5. van Riel PLCM, van Gestel AM, van de Putte LBA. Development and validation of response criteria in rheumatoid arthritis: steps towards an international consensus on prognostic markers. *Br J Rheumatol* 1996;35(Suppl. 2):4–7.
6. Boers M, Tugwell P. OMERACT conference questionnaire results. OMERACT Committee. *J Rheumatol* 1993;20:552–4.
7. Paulus HE, Egger JR, Ward JR, Williams HJ. Analysis of improvement in individual rheumatoid arthritis patients treated with disease-modifying antirheumatic drugs, based on the findings in patients treated with placebo. *Arthritis Rheum* 1990;33:477–84.
8. Tugwell P, Boers M, Baker P, Wells G, Snider J. Endpoints in rheumatoid arthritis. *J Rheumatol* 1994;21:2–8.
9. Boers M, Tugwell P, Felson DT *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.
10. van der Heijde DMFM, Van't Hof MA, van Riel PLCM *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.
11. van Gestel AM, Haagsma CJ, van Riel PLCM. Validation of rheumatoid arthritis improvement criteria that include simplified joint counts. *Arthritis Rheum* 1998;41:1845–50.
12. 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.
13. van Gestel AM, Prevoo MLL, Van't Hof MA, van Rijswijk MH, van de Putte LBA, van Riel PLCM. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. *Arthritis Rheum* 1996;39:34–40.
14. van Gestel AM, Anderson JJ, van Riel PLCM *et al.* ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. *J Rheumatol* 1999;26:705–11.
15. van der Heijde DMFM, Van't Hof MA, van Riel PLCM, van de Putte LBA. Validity of single variables and indices to measure disease activity in rheumatoid arthritis. *J Rheumatol* 1993;20:538–41.
16. Eberl G, Studnicka-Benke A, Hitzelhammer H, Gschnait F, Smolen JS. Development of a disease activity index for the assessment of reactive arthritis (DAREA). *Rheumatology* 2000;39:148–55.
17. Smolen JS, Kalden JR, Scott DL *et al.* European Leflunomide Study Group: Efficacy and safety of leflunomide compared with placebo and sulphasalazine in active rheumatoid arthritis: a double-blind, randomised, multicentre trial. *Lancet* 1999;353:259–66.
18. Strand V, Cohen S, Schiff M *et al.* Treatment of active rheumatoid arthritis with leflunomide compared to placebo and methotrexate. *Arch Intern Med* 1999;159:2542–50.
19. Emery P, Breedveld FC, Lemmel EM *et al.* A comparison of the efficacy and safety of leflunomide and methotrexate for the treatment of rheumatoid arthritis. *Rheumatology* 2000;39:655–65.
20. Arnett FC, Edworthy SM, Bloch DA *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
21. Steinbrocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. *J Am Med Assoc* 1949;140:659–62.
22. Wells GA, Tugwell P, Kraag GR, Baker PRA, Groh J, Redelmeier DA. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. *J Rheumatol* 1993;20:557–60.
23. Schiff MH, Strand V, Oed C, Loew-Friedrich I. Leflunomide: efficacy and safety in clinical trials for the treatment of rheumatoid arthritis. *Drugs Today* 2000;36:383–94.
24. Fuchs HA, Brooks RH, Callahan LF, Pincus T. A simplified twenty-eight-joint quantitative articular index in rheumatoid arthritis. *Arthritis Rheum* 1989;32:531–7.
25. Smolen JS, Breedveld FC, Eberl G *et al.* Validity and reliability of the twenty-eight-joint count for the assessment of rheumatoid arthritis activity. *Arthritis Rheum* 1995;38:38–43.

26. Wolfe F, Cathey MA. The assessment and prediction of functional disability in rheumatoid arthritis. *J Rheumatol* 1991;18:1298–306.
27. van der Heide A, Jacobs JW, Dinant HJ, Bijlsma JW. The impact of endpoint measures in rheumatoid arthritis clinical trials. *Semin Arthritis Rheum* 1992;21:287–94.
28. Anderson JJ, Felson DT, Meenan RF, Williams HJ. Which traditional measures should be used in rheumatoid arthritis clinical trials? *Arthritis Rheum* 1989;32:1093–9.
29. Bombardier C, Raboud J, Auranofin Cooperating Group. A comparison of health-related quality-of-life measures for rheumatoid arthritis research. *Control Clin Trials* 1991;12(Suppl.):243–56S.
30. Anderson JJ, Chernoff MC. Sensitivity to change of rheumatoid arthritis clinical trial outcome measures. *J Rheumatol* 1993;20:535–7.
31. Plant MJ, Williams AL, O'Sullivan MM, Lewis PA, Coles EC, Jessop JD. Relationship between time-integrated C-reactive protein levels and radiologic progression in patients with rheumatoid arthritis. *Arthritis Rheum* 2000;43:1473–7.
32. Aman S, Paimela L, Leirisalo-Repo M *et al*. Prediction of disease progression in early rheumatoid arthritis by ICTP, RF and CRP. A comparative 3-year follow-up study. *Rheumatology* 2000;39:1009–13.
33. Listing J, Rau R, Muller B *et al*. HLA-DRB1 genes, rheumatoid factor, and elevated C-reactive protein: independent risk factors of radiographic progression in early rheumatoid arthritis. Berlin Collaborating Rheumatological Study Group. *J Rheumatol* 2000;27:2100–9.
34. Drossaers-Bakker KW, De Buck M, van Zeben D, Zwinderman AH, Breedveld FC, Hazes JMW. Long-term course and outcome of functional capacity in rheumatoid arthritis: The effect of disease activity and radiologic damage over time. *Arthritis Rheum* 1999;42:1854–60.
35. Wolfe F, Hawley DJ. The long-term outcomes of rheumatoid arthritis: work disability: A prospective 18 year study of 823 patients. *J Rheumatol* 1998;25:2108–17.
36. Scott DL, Pugner K, Kaarela K *et al*. The links between joint damage and disability in rheumatoid arthritis. *Rheumatology* 2000;39:122–32.