EXTENDED REPORT

Performance of response criteria for assessing peripheral arthritis in patients with psoriatic arthritis: analysis of data from randomised controlled trials of two tumour necrosis factor inhibitors

J Fransen, C Antoni, P J Mease, W Uter, A Kavanaugh, J R Kalden, P L C M Van Riel

See end of article for authors' affiliations
....

Correspondence to:
J Fransen, Department of Rheumatology, Radboud University Nijmegen Medical Centre, PO Box 9101, NL-6500HB Nijmegen, The Netherlands;
j.fransen@reuma.umcn.nl

Background: In recent clinical trials in patients with psoriatic arthritis (PsA), the response criteria and disease activity measures that have been used were those developed for rheumatoid arthritis. However, these have not yet been validated in PsA.

Objective: To compare the responsiveness and discriminative capacity of the psoriatic arthritis response criteria (PsARC), American College of Rheumatology (ACR) and European League Against Rheumatism (EULAR) response criteria and the Disease Activity Score (DAS) and core-set measures in patients with PsA and peripheral arthritis, using the data from two randomised placebo-controlled trials of tumour necrosis factor inhibitors.

Methods: In an infliximab trial, 104 patients with active PsA were randomised to receive placebo or infliximab for 16 weeks. In an etanercept trial, 60 patients with active PsA were randomised to receive placebo or etanercept for 12 weeks. Data from baseline and the end of the intervention phase were used from each study. Responsiveness was assessed using the standardised response mean and effect size. Capacity to discriminate between the active drug and placebo was assessed using t values or a χ^2 test. Measures were ranked in order of their t value or χ^2 value.

Results: The EULAR criteria performed better in discriminating the active drug from placebo than the ACR20 improvement criteria, which in turn performed better than the PsARC. It was also found that the pooled indices (DAS and DAS28) were generally more responsive, and performed better in discriminating active drug from placebo, than the single core-set measures.

Conclusion: Response criteria and pooled indices developed for rheumatoid arthritis are useful for the assessment of arthritis in PsA clinical trials.

A growing awareness of the severity of psoriatic arthritis (PsA), as regards to the effect of the arthritis and other manifestations of the disease, has been observed. A greater understanding of the pathophysiology of PsA, combined with progress in biotechnology, has resulted in the development of new, highly effective agents for the treatment of PsA. Although the peripheral arthritis of PsA shares some clinical characteristics with that of rheumatoid arthritis, PsA shows some distinct features. Nevertheless, in assessing the response to treatment of peripheral arthritis, outcome measures have been largely borrowed from those developed for rheumatoid arthritis.[1–3] These include the individual rheumatoid arthritis core-set measures, the American College of Rheumatology (ACR) improvement criteria, the Disease Activity Score (DAS) and the European League Against Rheumatism (EULAR) response criteria.[4–9] These measures, as well as composite criteria suggested for PsA (commonly called psoriatic arthritis response criteria (PsARC)),[10 11] have never been validated in PsA. Therefore, we compared the responsiveness and discriminative capacity of those rheumatoid arthritis response criteria and activity measures in patients with PsA having peripheral arthritis, using the data from two randomised placebo-controlled trials of tumour necrosis factor (TNF) inhibitors.

METHODS

Data

The data were derived from two recently published phase II randomised placebo-controlled trials of TNF inhibitors in patients with PsA.[11 12] For the analysis, only data from baseline and the end of the intervention phase were used for each study.

In the trial by Antoni et al[12] (Infliximab Multinational Psoriatic Arthritis Controlled Trial (IMPACT)), 104 patients with active psoriatic arthritis (defined as ≥5 swollen joints and ≥5 tender or painful joints) were randomised to receive placebo or infliximab for 16 weeks.

In the trial by Mease et al[11] (etanercept), 60 patients with active psoriatic arthritis (defined as ≥3 swollen joints and ≥3 tender or painful joints) were randomised to receive placebo or etanercept 25 mg twice weekly, for 12 weeks. In both trials, there were no patients with exclusive distal inter-phalangeal (DIP) involvement.[13] According to definitions by Helliwell et al,[14] 95% of the patients had polyarthritis (>5 joints affected) at baseline. The relevant ethics committees approved the trials and all patients provided written informed consent.

....

Table 1 An overview of the response criteria

PsARC	ACR criteria	EULAR criteria
Response is defined by improvement in at least 2 of the 4 following measures, one of which must be joint swelling or tenderness, and no worsening in any of the 4 measures: PtGA of articular disease (1–5) and PhGA of articular disease (1–5): improvement = decrease by one category, worsening = increase by one category. Joint pain/tenderness score and joint swelling score: improvement = decrease by 30%, worsening = increase by 30%.	Improvement is defined by at least 20% improvement in TJC and in SJC, and at least 20% improvement in 3 of the 5 measures: ESR or CRP PhGA of disease activity PtGA of disease activity Patient assessment of pain Disability	A good response is defined as reaching a DAS $\leqslant 2.4$ or a DAS28 $\leqslant 3.2$ ("low" disease activity) in combination with an improvement >1.2 (twice the measurement error) in DAS or DAS28. A non-response is defined as an improvement $\leqslant 0.6$, and also as an improvement $\leqslant 1.2$ with a DAS>3.7 or DAS28>5.1 ("high" disease activity). All other possibilities are defined as a moderate response.

ACR, American College of Rheumatology[5]; CRP, C reactive protein; DAS, Disease Activity Score[7]; ESR, erythrocyte sedimentation rate; EULAR, European League Against Rheumatism[8]; PhGA, physician global assessment; PsARC, psoriatic arthritis response criteria[10]; PtGA, patient global assessment; SJC, swollen joint count; TJC, tender joint count.

Assessments and measures

The patient assessments, physical examinations and laboratory assessments were similar in both trials. Arthritis pain and patient global assessments (PtGA) of disease activity were rated using the Likert scale (0–5), as well as by a Visual Analogue Scale (VAS, 0–100) in the IMPACT trial. Disability was assessed using the Disability Index (0–3) of the Health Assessment Questionnaire (HAQ). Duration of morning stiffness was rated in minutes. Physician global assessment (PhGA) of disease activity was rated using the Likert scale (0–5), as well as by a VAS (0–100) in the IMPACT trial. A 76 swollen-joint count (SJC76, 0–76) and a 78 tender-joint count (TJC78, 0–78) were used in the etanercept trial, and a 66 SJC (SJC66, 0–66) and a 68 TJC (TJC68, 0–68) in the IMPACT trial. All counts included the DIP joints of the hands and feet. Reduced joint counts (TJC28 and SJC28) were calculated from the extended joint counts. The Ritchie Articular Index (RAI 0–75) was calculated using the weighted TJC, omitting the subtalar joints that were not assessed. The Psoriasis Area and Severity Index (PASI 0–72) was assessed to determine skin involvement. Erythrocyte sedimentation rate (mm/h) and CRP (mg/l) were determined as measures of the acute-phase response. We refer the readers to the original publications for more details.[11] [12]

Indices for arthritic activity that were calculated from the data were the DAS and the Disease Activity Score of 28 joint counts (DAS28).[6] [7] The DAS and the DAS28 were calculated as

$$DAS = 0.53938\sqrt{(RAI)}+0.06465(SJC44)+0.330\ln(ESR)+0.0072(PtGA)$$

and

$$DAS28 = 0.56\sqrt{(TJC28)}+0.28\sqrt{(SJC28)}+0.70\ln(ESR)+0.014(PtGA).[6] [7]$$

Appropriate formulas omitting the global assessment were used for the etanercept trial, because PtGA was not assessed using a VAS scale.[6] [7]

Improvement criteria (table 1) calculated were the EULAR response criteria,[8] [9] the ACR improvement criteria with 20%, 50% and 70% cut-off points,[5] and the PsARC.[10] [11] Full available joint counts were used for PsARC and ACR criteria, thus including DIP joints. The PsARC use PtGA and PhGA in a Likert-scale format (improvement = decrease by 1 category, worsening = increase by 1 category) and TJC and SJC (improvement = decrease by 30%, worsening = increase by 30%).[10] Accordingly, patients were classified as responder when there was improvement in at least two of the four measures, one of which must be a joint score, and no worsening in any of the four measures. Additional low-disease activity criteria calculated were criteria for minimal disease activity in rheumatoid arthritis as proposed by the Outcome Measures in Rheumatoid Arthritis Clinical Trials

(OMERACT) conference in 2004. They defined minimal disease activity as the absence of swollen or tender joints and an ESR $\leqslant 10$, or having low values for at least 5 of the 7 core-set measures: not more than 1 tender joint (out of 28), not more than 1 swollen joint (out of 28), pain (VAS) $\leqslant 20$ mm, PtGA (VAS) $\leqslant 20$ mm, PhGA (VAS) $\leqslant 15$ mm, ESR $\leqslant 20$ mm/h and HAQ $\leqslant 0.5$ points[15]; the low-disease activity criteria of the DAS ($\leqslant 2.4$) and the DAS28 ($\leqslant 3.2$),[8] [9] and "near-remission" low-disease activity by the DAS (<1.6) and the DAS28 (<2.6).[16] [17] In addition, a global statistical test (GST) was calculated from the seven available measures of the core set for rheumatoid arthritis randomised controlled trials.[18] A GST is sometimes advised for trials with multiple end points to increase the power. The GST was calculated by ranking the patients according to their response for every core-set measure separately, followed by addition of all seven rank numbers for every patient.[18] The resulting summed rank values were analysed using a t test. The GST approach is very powerful in randomised controlled trials, but it is a measure of change aimed solely at statistical testing; the summed rank values are not clinically interpretable and cannot be compared across randomised controlled trials.

Statistical analysis

Responsiveness of the continuous outcome measures was assessed using the standardised response mean (SRM, calculated as mean change/SD change) and effect size (calculated as mean change/SD baseline). SRM and effect size can be seen as signal-to-noise ratios. Within trials, higher values of SRM or effect size point to better sensitivity to change or responsiveness of a particular measure. It is not clear whether SRM or effect size is the better responsiveness statistic, although we prefer the SRM because it shows change in its nominator as well as its denominator. The discriminative capacity of the continuous outcome measures was tested using t values of the difference in change between the active drug and placebo. Larger t values point to better discrimination and lead to smaller p values for the difference between active drug and placebo. Next, all measures were ranked according to the t values, as the aim of the trials was the discrimination of the active drug from placebo.[19] The t value of the GST can be used as a reference, as a GST is supposed to represent maximal power when combining existing measures.[18] The discriminative capacity of the different response criteria and low-disease activity criteria was tested using the Mantel–Haenszel χ^2, allowing the comparison of dichotomous and trichotomous measures with each other. For these analyses, no attempt was made to pool the data of both trials, mainly because of the differences in drugs and timing of the follow up. SRMs, t values and χ^2 values are therefore best interpreted within a trial, not

Table 2 Responsiveness of individual items and indices ranked by t value in the Infliximab Multinational Psoriatic Arthritis Controlled Trial

	Active drug					Placebo					
	Baseline	Mean Δ	SE$_Δ$	SRM	ES	Baseline	Mean Δ	SE$_Δ$	SRM	ES	t Value
GST	–	460	16.2	3.96	–	–	226	14.9	2.15	–	10.6
DAS28 (0–10)	5.5	−2.6	0.2	−1.87	−2.30	5.4	−0.2	0.1	−0.21	−0.20	9.8
PhGA (0–100)	53.6	−33.9	3.2	−1.49	−1.84	52.4	0.3	2.4	0.02	0.01	8.5
DAS (0–10)	4.1	−1.9	0.2	−1.77	−2.13	3.9	−0.1	0.2	−0.13	−0.15	8.2
Pain (0–100)	53.4	−31.1	3.1	−1.40	−1.46	56.0	−0.1	2.8	−0.01	0.00	7.3
PtGA (0–100)	52.4	−27.6	3.2	−1.21	−1.18	56.0	1.0	3.1	0.05	0.05	6.4
TJC68 (0–68)	23.7	−13.1	1.6	−1.14	−0.96	20.4	0.8	1.7	0.07	0.07	6.0
ESR (mm/h)	32.7	−24.3	2.8	−1.24	−1.10	29.5	−4.6	1.9	−0.34	−0.24	5.7
TJC28 (0–28)	10.6	−6.2	0.7	−1.14	−0.98	9.7	−0.11	0.8	−0.02	−0.02	5.4
MS (min)	125.8	−91.5	21.9	−0.59	−0.59	149.9	−24.1	28.1	−0.12	−0.09	5.0
SJC66 (0–66)	14.6	−9.7	1.2	−1.15	−1.27	14.7	−1.0	1.3	−0.11	−0.14	5.0
HAQ (0–3)	1.15	−0.56	0.07	−1.18	−0.79	1.18	−0.01	0.04	−0.03	−0.01	4.8
SJC44 (0–44)	11.9	−7.7	1.0	−1.08	−1.26	11.7	−1.0	1.1	−0.13	−0.14	4.7
RAI (0–75)	13.5	−6.7	0.8	−1.19	−0.86	11.8	−0.2	1.3	−0.02	−0.02	4.3
SJC28 (0–28)	8.7	−5.7	0.9	−0.93	−1.13	9.1	−1.3	0.8	−0.21	−0.23	3.8
PASI (0–72)	5.1	−4.1	0.8	−0.82	−0.69	4.2	0.6	0.4	0.22	0.11	2.6
CRP (mg/l)	21.7	−11.3	3.4	−0.61	−0.57	31.1	−8.9	3.3	−0.40	−0.37	0.6

CRP, C reactive protein; DAS, Disease Activity Score; ES, effect size; ESR, erythrocyte sedimentation rate; GST, global statistical test; HAQ, Health Assessment Questionnaire; Mean Δ, mean change; MS, morning stiffness; PASI, Psoriasis Area and Severity Index; PhGA, physician global assessment; PtGA, patient global assessment; RAI, Ritchie Articular Index; SE$_Δ$, standard error of mean change; SJC, swollen joint count; SRM, standardised response mean; TJC, tender joint count.

across both trials. However, the relative ranking of measures was compared across the trials.

RESULTS
As indicated by the mean changes (tables 2 and 3), all measures improved in the active drug group in both trials. In the IMPACT trial, for example, the DAS28 showed a mean improvement of 2.6 points, the TJC68 a mean improvement of 13 joints and the HAQ a mean improvement of 0.56 points. In the etanercept trial, the DAS28 reduced by a mean of 1.7 points, the TJC68 improved by a mean of 12 joints and the HAQ improved by an average of 0.73 points. In both active drug groups the responsiveness statistics were >1.0 for most measures.

Placebo responses in terms of improvement, which are indicated by negative SRMs in tables 2 and 3, were largest in ESR, CRP, SJC and the DAS28 in the IMPACT trial, and in PhGA and the SJC in the etanercept trial.

All indices of disease activity performed better than the single variables in both trials with regard to responsiveness (SRM and effect size) and discriminating change in the placebo group from change in the active drug group (t value) than the single variables in both trials (tables 2 and 3). The SRM and effect size in the active drug group represent the responsiveness to change of a measure. The ability to discriminate between change in the active drug group and a probable change in the placebo group is given by the t values. A t value >1.96 is indicative of a significant (p<0.05) difference between the treatment groups. The clinical index that performed best, according to the ranking of the t value, was the DAS28, with a performance similar to the global statistical index GST. The single variable that performed best was the PhGA of disease activity. Patient ratings of pain and global disease activity were more responsive than joint counts in the IMPACT trial, irrespective of their format (0–5 or 0–100), but not in the etanercept trial. TJCs were more responsive than SJCs and the measures of the acute phase response. Further, extended joint counts were generally more responsive than reduced ones.

In both trials, the EULAR response criteria performed better ($χ^2$) than the ACR20 or PsARC criteria did in discriminating the active drug from placebo (table 4). The PsARC performed similar to or less well than the ACR20

criteria. Using data from the IMPACT trial, the ACR improvement criteria were recalculated using the more responsive VAS scales, instead of 0–5 rating scales, for global assessments and pain. The results of both methods of calculation were exactly the same (not shown).

The DAS28 low-disease activity criterion performed as well as the ACR20, PsARC and OMERACT criteria.

DISCUSSION
In this study, we showed that response criteria and pooled indices initially developed for rheumatoid arthritis are useful for assessing arthritis in PsA clinical trials. In this analysis we considered only measures of peripheral arthritis. Other important core domains of PsA that may be assessed in clinical trials include the involvement of the skin and nails, axial arthritis and the involvement of periarticular structures such as entheses.[3] The development and validation of useful outcome measures for these domains will facilitate more a complete definition of the effect of treatment, comparisons among therapeutic approaches and understanding of the course of PsA.[3] For assessing peripheral joint arthritis in PsA clinical trials, several measures that were primarily developed for use in rheumatoid arthritis are used. However, the potential value of these measures in PsA had not been shown. Therefore, we compared the responsiveness and discriminative capacity of existing response criteria and arthritic activity measures in two randomised placebo-controlled trials of TNF antagonists in PsA.[11 12] In the trials we studied, the ACR20 improvement criteria, the EULAR response criteria and the PsARC all showed significant differences between the active drug and placebo with p<0.001. When ranking the response measures according to their discriminative capacity, the EULAR criteria performed better in discriminating the active drug from placebo than the ACR20 improvement criteria, which in turn performed better than or as well as the PsARC.

This does not necessarily imply that the EULAR criteria will be "the best" response measure for all future PsA clinical trials. In rheumatoid arthritis, the discriminating capacity of ACR20 and EULAR criteria was reported to be similar in trials of conventional disease-modifying antirheumatic drugs.[20] In this study, we used clinical trials of highly effective biological agents. A factor contributing to the performance of the

Table 3 Responsiveness of individual items and indices, ranked by t value in the etanercept trial

	Active drug					Placebo					
	Baseline	Mean Δ	SE_Δ	SRM	ES	Baseline	Mean Δ	SE_Δ	SRM	ES	t Value
DAS28 (0–10)	5.2	−2.0	0.2	−2.22	−1.51	4.7	0.2	0.2	0.14	0.12	8.6
GST	–	190	9.2	3.76	–	–	91	8.3	2.10	–	7.9
DAS (0–10)	3.9	−1.7	0.2	−1.74	−1.55	3.9	−0.0	0.2	−0.01	−0.01	7.4
PhGA (0–5)	3.3	−2.1	0.2	−1.98	−2.84	3.4	−0.3	0.2	−0.34	−0.32	7.0
TJC78 (0–78)	22.7	−14.2	2.0	−1.32	−1.17	25.3	4.2	1.9	0.42	0.22	6.6
TJC68 (0–68)	19.8	−12.1	1.8	−1.25	−1.19	20.5	5.6	2.0	0.53	0.42	6.6
PtGA (0–5)	3.3	−2.0	0.2	−1.51	−1.99	2.9	−0.1	0.8	−0.18	−0.18	6.4
TJC28 (0–28)	7.7	−5.3	0.8	−1.17	−1.08	7.2	2.9	1.2	0.48	0.59	5.8
SJC66 (0–66)	14.2	−9.4	1.15	−1.53	−1.16	15.4	−1.1	1.0	−0.20	−0.10	5.4
MS (min)	56.5	−35.8	5.4	−1.21	−1.61	59.7	1.3	5	0.05	0.09	5.0
HAQ (0–3)	1.23	−0.73	0.10	−1.32	−1.28	1.18	−0.07	0.09	−0.16	−0.13	4.8
SJC44 (0–44)	10.6	−6.5	0.9	−1.35	−1.01	11.6	−0.9	4.0	−0.22	−0.11	4.7
RAI (0–75)	12.7	−8.0	1.3	−1.18	−1.13	14.7	0.8	1.4	0.10	0.08	4.6
SJC76 (0–76)	15.7	−10.6	1.3	−1.49	−1.13	18.3	−2.5	1.3	−0.37	−0.18	4.4
SJC28 (0–28)	6.8	−3.9	0.6	−1.27	−0.84	6.2	−0.7	0.5	−0.25	−0.15	4.1
CRP (mg/l)	22.5	−15.0	3.2	−0.86	−0.70	16.3	2.1	3.0	0.13	0.15	3.8
Pain (0–5)	2.9	−1.8	0.2	−1.36	−1.45	3.1	−0.1	0.2	−0.07	−0.07	3.7
ESR (mm/h)	26.6	−15.2	3.1	−0.91	−0.58	23.2	−0.7	3.2	−0.05	−0.03	3.2
PASI (0–72)	12.7	−5.0	1.3	−0.89	−0.62	7.8	−0.4	1.2	−0.08	−0.08	2.6

CRP, C reactive protein; DAS, Disease Activity Score; ES, effect size; ESR, erythrocyte sedimentation rate; GST, global statistical test; HAQ, Health Assessment Questionnaire; IMPACT, Infliximab Multinational Psoriatic Arthritis Controlled Trial; Mean Δ, mean change; MS, morning stiffness; PASI, Psoriasis Area and Severity Index; PhGA, physician global assessment; PtGA, patient global assessment; RAI, Ritchie Articular Index; SE_Δ, standard error of mean change; SJC, swollen joint count; SRM, standardised response mean; TJC, tender joint count.

Table 4 Discriminative capacity of response criteria

IMPACT Response criteria	Placebo (n = 52) % improved	Active drug (n = 52) % improved	χ^2_{MH}	p Value
EULAR28			53.4	<0.001
Good+moderate	26	92%		
Good	0	60		
EULAR			48.6	<0.001
Good+moderate	26	88		
Good	0	63		
ACR20	10	67	36.2	<0.001
ACR50	0	48	32.6	<0.001
ACR70	0	29	17.4	<0.001
PsARC	30	82	27.9	<0.001
ΔDAS28>1.2	13	83	49.4	<0.001
ΔDAS>1.2	10	67	36.2	<0.001
Low-disease activity criteria	**% reached**	**% reached**	χ^2_{MH}	**p Value**
DAS28 ⩽ 3.2	4	61	35.3	<0.001
DAS ⩽ 2.4	8	65	33.4	<0.001
OMERACT MDAS	6	50	25.1	<0.001
DAS28 ⩽ 2.6	2	41	21.2	<0.001
DAS ⩽ 1.6	0	35	20.0	<0.001
Etanercept Response criteria	**Placebo (n = 30) % improved**	**Active drug (n = 30) % improved**	χ^2_{MH}	**p Value**
EULAR28			26.2	<0.001
Good+moderate	15	93		
Good	8	59		
EULAR			24.1	<0.001
Good+moderate	19	89		
Good	4	44		
ACR20	15	73	19.6	<0.001
ACR50	4	50	14.8	0.001
ACR70	0	13	3.8	0.05
PsARC	33	90	19.3	<0.001
ΔDAS28>1.2	7	80	32.2	<0.001
ΔDAS>1.2	7	60	18.9	<0.001
Low-disease activity criteria	**% reached**	**% reached**	χ^2_{MH}	**p Value**
DAS28 ⩽ 3.2	11	69	19.0	<0.001
DAS ⩽ 2.4	15	69	16.4	<0.001
OMERACT MDAS	7	63	20.8	<0.001
DAS28 ⩽ 2.6	11	59	13.5	0.002
DAS ⩽ 1.6	4	45	12.4	0.004

ACR, American College of Rheumatology; DAS, Disease Activity Score; EULAR, European League Against Rheumatism; IMPACT, Infliximab Multinational Psoriatic Arthritis Controlled Trial; MDAS, minimal disease activity in rheumatoid arthritis; OMERACT, Outcome Measures in Rheumatoid Arthritis Clinical Trials; PsARC, psoriatic arthritis response criteria; χ^2_{MH}, χ^2 Mantel Haenszel.

EULAR criteria in these trials is that many active drug-treated patients reached the low-disease activity criterion of the EULAR response criteria, but few placebo-treated patients did. Although the ACR50 and ACR70 criteria have been proposed as meaningful end points when testing very effective treatments, these stringent measures were less discriminative than the ACR20 criteria in this study. This is consistent with findings in rheumatoid arthritis.[21] The low-disease activity criteria that were tested in this study (eg, DAS28<3.2) discriminated well between the active drug and placebo. However, as the cut-off point was lowered (eg, DAS28<2.6 or OMERACT MDAS), smaller numbers of active drug-treated patients fulfilled these criteria and the discrimination was less clear.

The pooled indices (GST, DAS, DAS28) were generally more responsive and discriminating than single measures in the trials we studied. The DAS28 performed better than the DAS; this was unexpected because the DAS includes a larger number of counted joints. A reason for the lower discriminative ability of the DAS as compared with the DAS28 may be that in the DAS relatively much weight is given to the RAI, which in both trials did not perform as well as the ungraded TJCs, including the 28TJC. The discriminative capacity of the RAI was also less than the TJC in a study comparing responsiveness of outcome measures in early rheumatoid arthritis.[19] The RAI may be less sensitive to change because of observer error due to the grading of tenderness, as well as the scoring of the proximal interphalangeal joints of each hand, metacarpophalangeal joints of each hand, and metatarsophalangeal joints of each foot as a single unit. The 28SJC and the 28TJC showed less discriminative ability than the more extended joint counts. However, their combination in the DAS28 seemed advantageous, as the discriminative ability of the DAS28 was close to the performance of the Global Statistical Index, which included the full joint counts. The reason for not losing all discriminative information when omitting the feet and DIP joints from a full joint count is that changes in extended and reduced joint counts are related. Certainly, in individual patients, excluding the feet and DIP joints from joint counts may lead to underestimation of disease activity. This is especially so in patients with monoarthritis or oligoarthritis and in patients with predominantly, or exclusively, DIP involvement. Underestimation of disease activity by omission of joints is not a major issue in larger randomised controlled trials because this underestimation will occur at random in the arms of the trial. Most of the patients in the trials we studied had involvement of at least five joints (tender or swollen), so we could not study the response in patients with oligoarticular disease. In those patients, it may be more difficult to reach and measure response.

The results of our study apply only to clinical trials. In practice, the goal of treatment is principally to reach low-disease activity, and therefore response criteria are generally less useful as they present change in status, whereas the interest is in knowledge about absolute status. The DAS or DAS28 are measures with absolute value and could therefore be interesting for assessing PsA arthritic activity in practice, but they present a level of disease activity ''as would the patient have RA''. Use of the DAS and DAS28 for patients with PsA in practice might be more appropriate in rheumatoid arthritis-like polyarthritis, but it is still not clear what a certain level of DAS or DAS28 means in PsA. Especially, there is no clarity on whether it is necessary to include DIP joints in the joint counts, even when isolated DIP involvement is seldom seen,[22 23] and whether the weights for the joint counts and ESR and PtGA should be changed or whether ESR can be omitted.

It can be concluded that response criteria and pooled indices developed for rheumatoid arthritis may be used in PsA clinical trials as well, at least in the case of active polyarticular presentation as predominantly expressed in trials assessing highly effective TNF inhibitors. Further research may help to define the extent to which development of measures specific for PsA might perform better.

....

Authors' affiliations
J Fransen, P L C M Van Riel, Department of Rheumatology, University Medical Centre St Radboud, Nijmegen, The Netherlands
C Antoni, J R Kalden, Department of Medicine III, Friedrich-Alexander University, Erlangen, Germany
P J Mease, Seattle Rheumatology Associates, Swedish Medical Center, Seattle, Washington, USA
W Uter, Department of Medical Informatics, Friedrich-Alexander University, Erlanger, Germany
A Kavanaugh, Division of Rheumatology, Allergy, and Immunology, University of California, San Diego, California, USA

REFERENCES
1 **Gladman DD**, Helliwell P, Mease PJ, Nash P, Ritchlin C, Taylor W. Assessment of patients with psoriatic arthritis. A review of currently available measures. *Arthritis Rheum* 2004;**50**:24–35.
2 **Kavanaugh A**, Cassell S. Outcome measures in psoriatic arthritis. *Curr Rheumatol Rep* 2005;**7**:195–200.
3 **Mease PJ**, Antoni CE, Gladman DD, Taylor WJ. Psoriatic arthritis assessment tools in clinical trials. *Ann Rheum Dis* 2005;**64**(Suppl 2):49–54.
4 **Felson DT**, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, 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.
5 **Felson DT**, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:727–35.
6 **Van der Heijde DMFM**, Van't Hoff MA, Van Riel PLCM, Theunisse LAM, Lubberts EW, Van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. *Ann Rheum Dis* 1990;**49**:916–20.
7 **Prevoo MLL**, Van't Hoff MA, Kuper HH, Van Leeuwen MA, Van de Putte LBA, Van Riel PLCM. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 1995;**38**:44–8.
8 **Van Gestel AM**, Prevoo MLL, Van't Hoff MA, Van Rijswijk MH, Van de Putte LBA, et al. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. *Arthritis Rheum* 1996;**39**:34–40.
9 **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.
10 **Clegg DO**, Reda DJ, Mejias E, Cannon GW, Weisman MH, Taylor T, et al. Comparison of sulfasalazine and placebo in the treatment of psoriatic arthritis. *Arthritis Rheum* 1996;**39**:2013–20.
11 **Mease PJ**, Goffe BS, Metz J, VanderStoep A, Finck B, Burge DJ. Etanercept in the treatment of psoriatic arthritis and psoriasis: a randomised trial. *Lancet* 2000;**356**:385–90.
12 **Antoni CE**, Kavanaugh A, Kirkham B, Tutuncu Z, Burmester G, Schneider U, et al. Sustained benefits of infliximab therapy for dermatologic and articular manifestations of psoriatic arthritis: results from the infliximab multinational psoriatic arthritis controlled trial (IMPACT). *Arthritis Rheum* 2005;**52**:1227–36.
13 **Moll JMH**, Wright V. Psoriatic arthritis. *Semin Arthritis Rheum* 1973;**3**:55–78.
14 **Helliwell P**, Marchesoni A, Peters M, Barker M, Wright V. A re-evaluation of the osteoarticular manifestations of psoriasis. *Br J Rheumatol* 1991;**30**:339–45.
15 **Wells GA**, Boers M, Shea B, Brooks PM, Simon LS, Strand CV, et al. Minimal disease activity for rheumatoid arthritis: a preliminary definition. *J Rheumatol* 2005;**32**:2016–24.
16 **Prevoo MLL**, Van Gestel AM, Van't Hof MA, Van Rijswijk MH, Van de Putte LBA, Van Riel PLCM. Remission in a prospective study of patients with rheumatoid arthritis. *Br J Rheumatol* 1996;**35**:1101–5.
17 **Fransen J**, Creemers MC, Van Riel PL. Remission in rheumatoid arthritis: agreement of the disease activity score (DAS28) with the ARA preliminary remission criteria. *Rheumatology (Oxford)* 2004;**43**:1252–5.
18 **O'Brien PC**, Shampo MA. Statistical considerations for performing multiple tests in a single experiment. *Mayo Clin Proc* 1988;**63**:1140–3.

19 **Verhoeven AC**, Boers M, Van der Linden S. Responsiveness of the core set, response criteria, and utilities in early rheumatoid arthritis. *Ann Rheum Dis* 2000;**39**:966–74.

20 **Van Gestel AM**, Anderson JJ, Van Riel PLCM, Boers M, Haagsma CJ, Rich B, et al. ACR and EULAR improvement criteria have comparable validity in rheumatoid arthritis trials. *J Rheumatol* 1999;**26**:705–11.

21 **Felson DT**, Anderson JJ, Lange MLM, LaValley, MP. Should improvement in rheumatoid arthritis clinical trials be defined as fifty percent or seventy percent

improvement in core set measures, rather then twenty percent? *Arthritis Rheum* 1998;**41**:1564–70.

22 **Bruce IN**. Psoriatic arthritis: clinical features. In: Hochberg MC, et al. *Rheumatology*. Vol 2. 3rd edn. Toronto: Mosby, 2003.

23 **Alonso JCT**, Perez AR, Castrillo JMA, Garcia JB, Noriega JLR, Larrea CL. Psoriatic arthritis (PA): a clinical immunological and radiological study of 180 patients. *Br J Rheumatol* 1991;**30**: 245–50.

Clinical Evidence—Call for contributors

Clinical Evidence is a regularly updated evidence-based journal available worldwide both as a paper version and on the internet. *Clinical Evidence* needs to recruit a number of new contributors. Contributors are healthcare professionals or epidemiologists with experience in evidence-based medicine and the ability to write in a concise and structured way.

Areas for which we are currently seeking contributors:

- Pregnancy and childbirth
- Endocrine disorders
- Palliative care
- Tropical diseases

We are also looking for contributors for existing topics. For full details on what these topics are please visit www.clinicalevidence.com/ceweb/contribute/index.jsp
However, we are always looking for others, so do not let this list discourage you.

Being a contributor involves:

- Selecting from a validated, screened search (performed by in-house Information Specialists) epidemiologically sound studies for inclusion.
- Documenting your decisions about which studies to include on an inclusion and exclusion form, which we keep on file.
- Writing the text to a highly structured template (about 1500-3000 words), using evidence from the final studies chosen, within 8-10 weeks of receiving the literature search.
- Working with *Clinical Evidence* editors to ensure that the final text meets epidemiological and style standards.
- Updating the text every 12 months using any new, sound evidence that becomes available. The *Clinical Evidence* in-house team will conduct the searches for contributors; your task is simply to filter out high quality studies and incorporate them in the existing text.

If you would like to become a contributor for *Clinical Evidence* or require more information about what this involves please send your contact details and a copy of your CV, clearly stating the clinical area you are interested in, to CECommissioning@bmjgroup.com.

Call for peer reviewers

Clinical Evidence also needs to recruit a number of new peer reviewers specifically with an interest in the clinical areas stated above, and also others related to general practice. Peer reviewers are healthcare professionals or epidemiologists with experience in evidence-based medicine. As a peer reviewer you would be asked for your views on the clinical relevance, validity, and accessibility of specific topics within the journal, and their usefulness to the intended audience (international generalists and healthcare professionals, possibly with limited statistical knowledge). Topics are usually 1500-3000 words in length and we would ask you to review between 2-5 topics per year. The peer review process takes place throughout the year, and out turnaround time for each review is ideally 10-14 days.
If you are interested in becoming a peer reviewer for *Clinical Evidence*, please complete the peer review questionnaire at www.clinicalevidence.com/ceweb/contribute/peerreviewer.jsp