An Index of Patient Reported Outcomes (PRO-Index) Discriminates Effectively Between Active and Control Treatment in 4 Clinical Trials of Adalimumab in Rheumatoid Arthritis

THEODORE PINCUS, CECILIA CHUNG, OSCAR G. SEGURADO, INGRID AMARA, and GARY G. KOCH

ABSTRACT. ***Objective.*** To analyze 2 indices composed of the 3 patient reported outcomes (PRO) in the American College of Rheumatology (ACR) Core Data Set — physical function, pain, and global estimate — without joint count or laboratory data, for capacities to distinguish active from control treatments in 4 pivotal clinical trials.

Methods. Data from 4 clinical trials involving adalimumab, in combination with methotrexate or other disease-modifying antirheumatic drugs (DMARD) or as monotherapy, versus control treatment were made available to analyze properties of various indices. A categorical PRO-Index M was defined as "majority" improvement in 2 of the 3 PRO measures at 20%, 50%, and 70% levels; results were evaluated to analyze agreement with ACR20, ACR50, ACR70 responses and an "all Core Data Set measures" index based on 4 of the 7 measures having such levels of improvement. A continuous PRO-Index C was defined as the median or 2nd highest of 3 percentage differences from baseline to endpoint; results were evaluated to analyze agreement with a continuous ACR-N, "all Core Data Set measures" index, and Disease Activity Score 28 (DAS28).

Results. All indices distinguished active versus control treatment at similar levels, including PRO-Index M versus ACR20, ACR50, and ACR70 responses, and PRO-Index C versus DAS28.

Conclusion. PRO indices based only on patient questionnaire data, without joint counts or laboratory tests, may be useful quantitative measures of therapeutic efficacy for use in standard rheumatology clinical care. (First Release Oct 15 2006; J Rheumatol 2006;33:2146–52)

Key Indexing Terms:
QUESTIONNAIRES RHEUMATOID ARTHRITIS RANDOMIZED CONTROLLED TRIAL
ADALIMUMAB PATIENT INDEX

The American College of Rheumatology (ACR) Core Data Set[1-3] and Disease Activity Score 28 (DAS28)[4,5] are major advances to standardize measurement in rheumatoid arthritis (RA) clinical trials and other clinical research. Criteria based on 20%, 50%, and 70% improvement in ACR Core Data Set measures (ACR20, 50, 70) have been developed for clinical trials[6]. The ACR response criteria represent change scores from baseline to endpoint, and cannot be used as an absolute clinical measure. DAS scores have an absolute value rather than being based on change, and are useful in both clinical trials and standard clinical care.

From the Vanderbilt University Medical Center, Nashville, TN; Abbott Laboratories, Abbott Park, IL; Quintiles, Inc., Durham, NC; and the University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

T. Pincus, MD, Professor of Medicine; C. Chung, MD, PhD, Vanderbilt University Medical Center; O.G. Segurado, MD, PhD, Abbott Laboratories; I. Amara, DrPH, Quintiles, Inc.; G.G. Koch, PhD, University of North Carolina at Chapel Hill.

Address reprint requests to Dr. T. Pincus, Division of Rheumatology and Immunology, Vanderbilt University School of Medicine, 203 Oxford House, Box 5, Nashville, TN 37232-4500. E-mail:t.pincus@vanderbilt.edu

Accepted for publication June 6, 2006.

These indices are advances used primarily to assess RA in clinical trials and clinical research. Most standard rheumatology clinical care, however, is conducted according to empirical qualitative impressions rather than quantitative clinical measurement. Formal quantitative joint counts, although regarded by rheumatologists as the most important means to assess RA[5], and needed to calculate ACR criteria or a DAS, generally are not performed at most visits of patients with RA to most rheumatologists[6]. The only quantitative measure recorded at most visits is a laboratory test such as erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP), which may be normal in 40% of patients[7]. Therefore, ACR criteria or a DAS generally are not available in standard rheumatology care, despite an excellent website (www.das-score.nl) and calculators that simplify computation of the DAS.

A pragmatic quantitative measure that does not require joint counts or laboratory tests, for which the patient does almost all the work, might be of value to rheumatologists and their patients in standard clinical care. Three of the 7 components of the ACR Core Data Set, physical function, pain, and global estimate, are patient reported outcome (PRO) measures. A quantitative index of only these 3 patient self-report

questionnaire measures could facilitate quantitative assessment in standard clinical care.

We have reported that an index of patient measures yielded results similar to standard indices to distinguish results of treatment with leflunomide or methotrexate (MTX) versus placebo[8,9]. A "majority" PRO-Index M, similar in design to ACR20, 50, 70[4], was comparable to ACR response criteria; "continuous" PRO-Index C, similar in design to a continuous ACR-N[8], was similar to the ACR-N and DAS28 in the same clinical trial[9]. These results were all based on a single clinical trial. A PRO-Index termed the Patient Activity Score has been reported to be in agreement with other measures in standard clinical care[11].

Further analyses of additional RA clinical trials involving different therapeutic regimens appear desirable to assess the generalizability of PRO-Indices to distinguish between active versus control therapies at levels similar to ACR20 and DAS28. Therefore, we analyzed PRO-Index M and PRO-Index C for capacity to distinguish active from control treatments in 4 additional RA clinical trials involving adalimumab versus control treatments.

MATERIALS AND METHODS

Clinical trials. Clinical data from the adalimumab and control arms of 4 clinical trials were analyzed: ARMADA[12] and DE019[13], in which a regimen of adalimumab plus MTX was compared to placebo plus MTX; DE011[14], in which adalimumab monotherapy was compared to placebo; and STAR[15], in which a regimen of adalimumab plus other disease-modifying antirheumatic drugs (DMARD) was compared to placebo plus DMARD. Adalimumab treatments included study arms with the standard 40 mg every-other-week dose as well as other dosages. Analyses of all 4 trials were performed at 24 weeks, or at last observation during the 24 week-time period with sufficient data for determination of the indices, so the results would be based on similar periods, although 2 trials were conducted over longer periods.

A dichotomous, categorical PRO-Index M was computed from scores for physical function, pain, and patient's global estimate, i.e., the 3 PRO measures on a patient questionnaire in the ACR Core Data Set. This "majority" method is based on changes in the individual ACR Core Data Set measures from baseline to endpoint. Responders are defined as those patients who meet the required criteria of at least 20%, 50%, or 70% improvement, respectively, for at least 2 of the available measures. The index was computed only for patients who had available 2 of 3 measures among physical function, pain, or patient's global estimate at week 24, or last observation prior to week 24.

The primary evaluation of PRO-Index M was with respect to standard ACR20, ACR50, and ACR70 responses[4]. PRO-Index M also was evaluated relative to a categorical "all ACR Core Data Set" "majority" index, computed from all 7 measures, i.e., the 3 PRO measures, plus swollen joint count, tender joint count, assessor's global estimate, and C-reactive protein (CRP), without specific prioritization for changes in joint counts. Responders were defined as those patients who met the required criteria of at least 20%, 50%, or 70% improvement for at least 4 of the available measures. The index was computed only for patients who had available at least 4 of the 7 ACR Core Data Set measures at week 24, or at a last observation prior to week 24. These indices are analogous to ACR20, ACR50, or ACR70 responses[4] for which a "majority" of measures are required to show improvement at various levels.

A continuous version of the PRO index (PRO-Index C) was computed as the second highest of the 3 PRO measures for percentage change, (median value if all 3 measures were available, or the smallest of 2 available measures if one was missing)[9]. DAS28 was computed for patients who had all necessary components available at the last visit: tender joints, swollen joints, CRP, and patient's global estimate[16]. DAS28 was analyzed not only directly for its absolute change from baseline to endpoint, but also as percentage change, so as to have analyses comparable to the other continuous indices. Two additional continuous indices were analyzed: (a) A "continuous" version of the "all ACR Core Data Set" index is based on the median value for percentage change of all 7 ACR Core Data Set measures, or the fourth largest of the available measures if one or more was missing[9]. (b) A continuous ACR-N[8] was computed as the lowest of the percentage changes for tender joints, swollen joints, and the median (i.e., third largest) for the other 5 measures, or the second largest if one or more was missing.

Analyses according to both categorical "majority" type and continuous indices were also performed according to 4 additional indices: (a) PRO-Index M plus CRP; (b) PRO-Index C plus CRP; (c) "assessor-only index" involving 2 of 3 assessor measures: swollen joint count, tender joint count, assessor estimate of status; and (d) "assessor only plus CRP" involving 3 assessor-derived measures plus CRP[8,9]. The results were quite similar to ACR20, DAS28, PRO-Index M, and PRO-Index C, and are not presented here, but are available from the authors.

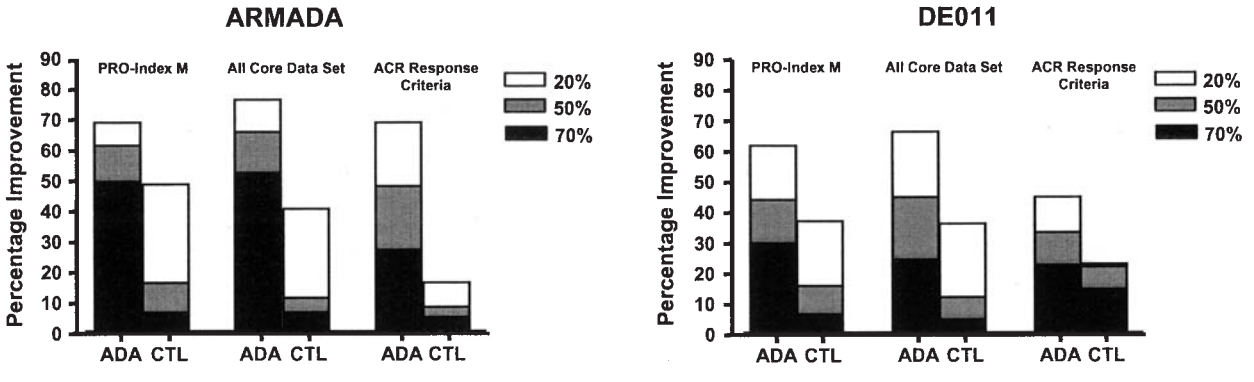
Statistical methods. Standard ACR20, ACR50, and ACR70 responses were computed as described[4]. The "all ACR Core Data Set" index is reported as the percentages of patients meeting criteria of at least 20%, 50%, and 70% improvements in each of the 4 clinical trials. Comparisons between treatments were made for each of the categorical indices from each study; statistical significance was tested according to the Fisher's exact test. Cross-tabulations were performed to evaluate possible agreement of ACR20 and ACR50 responses with PRO-Index 20% and PRO-Index 50% responses in individual patients.

The continuous indices, including PRO-Index C, a continuous "all ACR Core Data Set" index, ACR-N, and the DAS28 percentage improvement are reported as adjacent box plots for the active and control treatments. The tops and bottoms of boxes in the box plots indicate the 25th and 75th percentiles. Moreover, changes in shading of the boxes indicate the medians, dots in the boxes indicate the means, and lines below and above the boxes represent the extent of variability beyond the 25th and 75th percentiles (in terms of 150% of the distance between the 25th and 75th percentiles or the distance to the minimum or maximum values, whichever is smaller). For each study, comparisons between active and control treatments were made for the PRO-Index C, the "all ACR Core Data Set" index, ACR-N, and the change of DAS28 from baseline to endpoint in its own right, using 2-sample t tests (those for DAS28 were adjusted for baseline by analysis of covariance).

The PRO-Index C also has its entire distribution described graphically for each study according to cumulative probability curves. Values indicated by the vertical axis and related point on the curve reveal the percentage of patients who have the percentage change indicated on the horizontal axis. Each point indicates the percentage of patients with at least 2 of 3 patient measures showing a percentage improvement that is at least as large as the corresponding value on the horizontal axis in a manner analogous to a Kaplan-Meier cumulative survival curve. Negative values indicate worsening, while positive values indicate improvement.

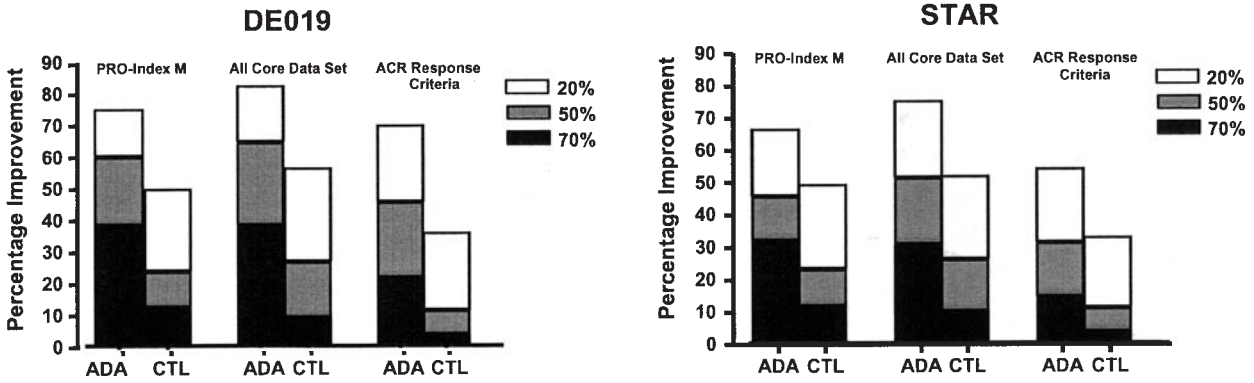
RESULTS

Categorical indices. The "PRO-Index M," the "majority" categorical index, "all ACR Core Data Set" index, and categorical ACR response criteria all indicated statistically significant advantages to active versus control treatment in all 4 clinical trials ($p < 0.01$ for all comparisons, except $p = 0.021$ for 20% improvement in ARMADA for the PRO-Index M; Figure 1). Differences between active versus control treatment were greater according to ACR criteria than according to the PRO-Index M for 3 of 4 comparisons of 20% improvement, but showed the opposite pattern for 8 of 8 comparisons of 50% and 70% improvement (Figure 1). Similarly, "all ACR Core Data Set" indicated greater discrimination of active versus



Improvement level	Estimated difference ADA vs. CTL*		
20%	20%	36%	53%
50%	45%	54%	40%
70%	43%	46%	22%

Improvement level	Estimated difference ADA vs. CTL*		
20%	25%	30%	22%
50%	28%	33%	14%
70%	23%	20%	11%



Improvement level	Estimated difference ADA vs. CTL*		
20%	25%	26%	34%
50%	36%	38%	34%
70%	26%	29%	18%

Improvement level	Estimated difference ADA vs. CTL*		
20%	17%	24%	21%
50%	22%	25%	21%
70%	20%	21%	11%

Figure 1. Analyses of 20%, 50%, and 70% changes according to 3 categorical indices, patient reported outcomes majority (PRO-Index M), "all ACR Core Data Set," and standard ACR improvement criteria. The PRO-Index M includes 3 patient questionnaire measures for physical function, pain, and patient global assessment. "All ACR Core Data Set" includes all 7 ACR Core Data Set measures without prioritization for joint counts (standard ACR improvement criteria require improvement for tender joints and swollen joints, as well as 3 of the other 5 criteria). The top of the black bar identifies the percentage with 70% or more improvement, the top of the grey bar cumulatively identifies the percentage with 50% or more improvement, and the top of the white bar cumulatively identifies the percentage with 20% or more improvement. The p values for all comparisons by Fisher exact test were < 0.001, other than p = 0.021 for 20% improvement according to the PRO-Index M in ARMADA, and p = 0.005 for 50%, and p = 0.001 for 70% improvement according to ACR improvement criteria in DE011. *All differences had p values < 0.001, except for p 0.01 for 50% improvement level for the ACR Index in DE011 and p < 0.05 for the 20% improvement level for PRO-M in ARMADA. ADA: adalimumab, CTL: control.

control for 2 of 4 comparisons at 20% improvement, and all 8 comparisons at 50% and 70% improvement.

Agreement was seen between ACR20 and ACR50 responses versus PRO-Index M 20% and 50% responses for 80% of patients, including patients treated with either adalimumab or control therapies (Table 1). A higher proportion of patients had 20% or 50% PRO-Index M responses in the absence of ACR20 or ACR50 responses (12–17%) compared to the opposite pattern of ACR20 responses in the absence of PRO-Index M 20% or PRO-Index M 50% responses (1–4%), due in large part to the greater stringency of ACR response criteria in requiring 20% improvement in tender and swollen joints.

The data suggest that prioritizing swollen and tender joint counts as requirements for ACR improvement criteria does not enhance sensitivity to detect differences between active and control treatments. However, this prioritization may be desirable because of added specificity. We interpret the findings to suggest that the PRO-Index M has capacities to distinguish active from control treatment that are similar to ACR improvement criteria[4] and "all ACR Core Data Set" measures, as might be expected since the PRO-Index M variables contribute importantly to ACR criteria.

Continuous indices. Statistically significant advantages to active versus control treatments (p < 0.001) were seen for PRO-Index C, ACR-N, all Core Data Set, and DAS28 (Figure 2). The mean percentage differences between adalimumab

Table 1. ACR 20 and ACR 50 responses and Pro-Index M 20 and 50 responses in 4 adalimumab clinical trials ARMADA, DE019, DE011, and STAR.

ACR 20 Responses	PRO-Index M 20% Responses	All Patients, no. (%)	Adalimumab Patients, no. (%)	Control Patients, no. (%)
+	+	580 (41.9)	384 (55.1)	196 (28.5)
+	–	41 (3.0)	27 (3.9)	14 (2.0)
–	+	220 (15.9)	91 (13.1)	129 (18.8)
–	–	543 (39.2)	195 (28.0)	348 (50.7)
ACR 50 Responses	PRO-Index M 50% Responses			
+	+	300 (21.7)	237 (34.0)	63 (9.2)
+	–	20 (1.4)	12 (1.7)	8 (1.2)
–	+	202 (14.6)	118 (16.9)	84 (12.2)
–	–	862 (62.2)	330 (47.3)	532 (77.4)
Total		1384 (100)	697 (100)	687 (100)

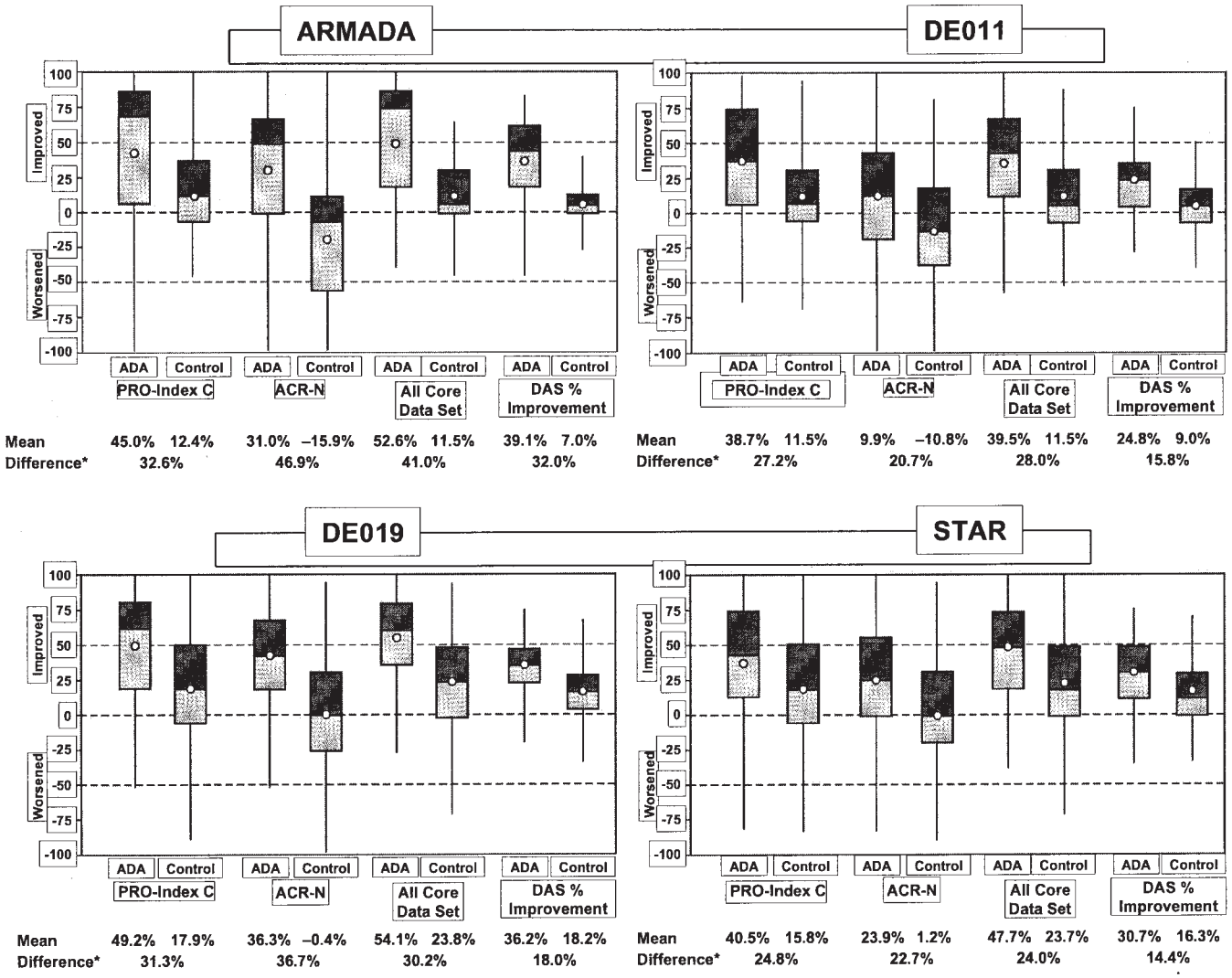


Figure 2. Box plots indicating percentages of patients who improved with active adalimumab versus control treatments according to 4 continuous indices: patient reported outcomes continuous (PRO-Index C), all ACR Core Data Set, ACR-N, and DAS 28. The tops and bottoms of the boxes in the box plots indicate the 25th and 75th percentiles, the locations of change in shading indicate the medians, the dots in the boxes indicate the means, and the lines below and above the boxes represent the extent of variability beyond the 25th and 75th percentiles. The p values for all treatment group comparisons according to analysis of variance t tests were < 0.001. *All differences had p values ≤ 0.001. ADA: adalimumab, CTL: control.

versus control groups according to the PRO-Index C was 33% in ARMADA, 27% in DE011, 31% in DE019, and 25% in STAR. The range of 25–33% was similar to the 14–32% differences in DAS28 in the 4 trials: 32% in ARMADA, 16% in DE011, 18% in DE019, and 14% in STAR. Results according to the other 2 continuous indices, ACR-N and All Core Data Set, were also similar to results according to the DAS28 (Figure 2). Mean improvements with active and control treatments were lowest according to ACR-N, including a higher proportion of patients in the control group with ACR-N values of 0 or less, indicating no improvement or worsening of status.

Probability plots illustrate in greater detail the capacity of the PRO-Index C to detect no improvement or worsening in each of the 4 clinical trials (Figure 3). Each point indicates the percentage of patients with at least 2 of 3 patient measures showing a percentage improvement at least as large as the corresponding value on the horizontal axis, analogous to a Kaplan-Meier cumulative survival curve, with negative values indicating worsening. No improvement was seen for about 15–25% of patients treated with adalimumab, compared with 35–40% of control patients. The lowest rate of control responses was seen in DE011, in which adalimumab was compared to placebo, in contrast to the other clinical trials that included MTX in the control group.

DISCUSSION

Our data indicate that PRO-Index M distinguishes adalimumab from control treatment groups in 4 clinical trials at levels similar to the categorical ACR20 index. Similarly, the PRO-Index C distinguishes the 2 treatment groups at levels comparable to the continuous DAS28, as well as ACR-N and other continuous indices developed in this study from the ACR Core Data Set. These findings extend previous evidence from one clinical trial of leflunomide, MTX, and placebo to 4 additional clinical trials using the anti-tumor necrosis factor biological agent adalimumab, highlighting the potential clinical value of PRO indices[8,9].

The ACR response criteria and DAS28 include joint counts, providing greater specificity for RA than PRO indices or assessor global measures. Therefore, ACR20, 50, and 70, and DAS28 appear desirable endpoints for clinical trials. However, PRO indices, without formal joint counts, appear of value to monitor drug therapy for RA in standard clinical care, with

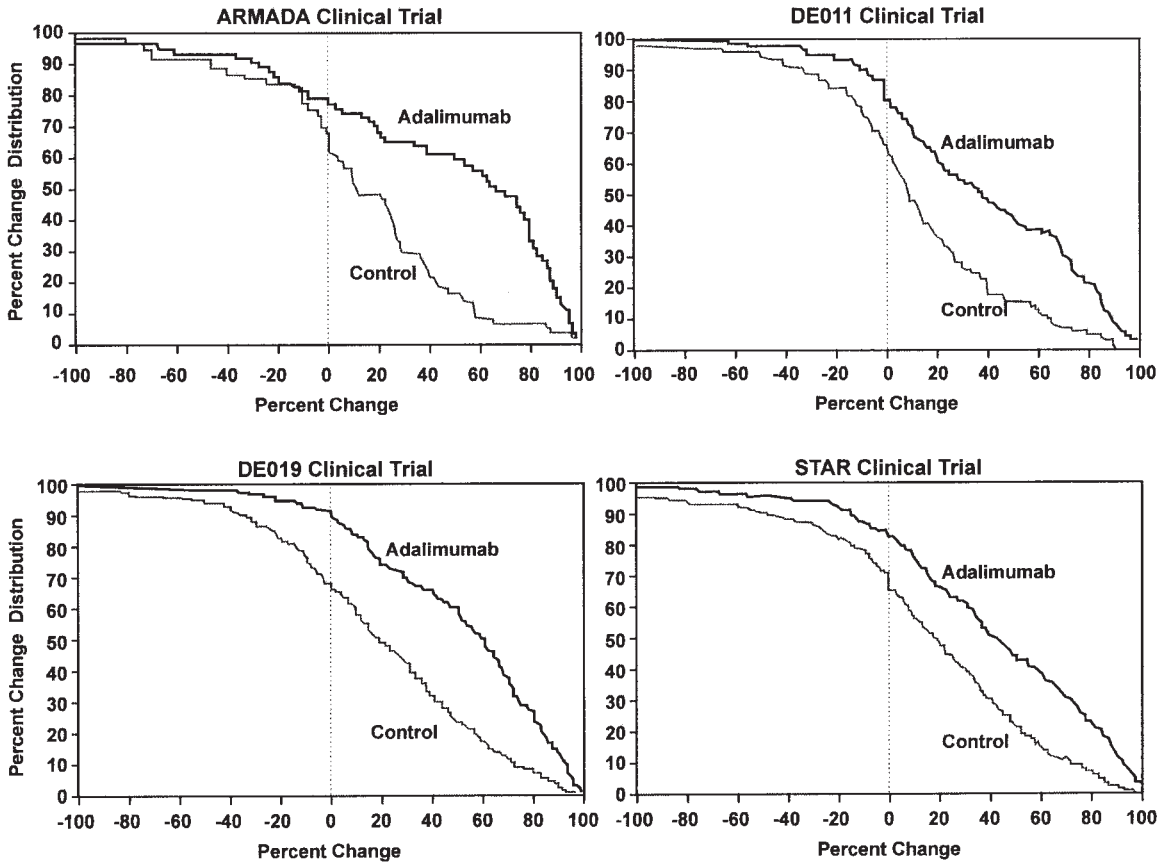


Figure 3. Cumulative probability curves for a patient reported outcomes continuous index (PRO-Index C), in 4 clinical trials comparing adalimumab versus control responses. A point on a curve and the vertical axis reveal the percentages of patients who have at least 2 of 3 patient measures with percentage improvement at least as large as the corresponding value on the horizontal axis in a manner analogous to a Kaplan-Meier cumulative survival curve. These displays include negative responses, indicating worsening clinical status.

little loss of information from indices that require formal joint counts.

Data from patient questionnaires are correlated significantly with data from traditional joint counts, radiographs, and laboratory tests[17], and can be more explanatory of other clinical information than these other measures[17]. Measures such as rheumatoid factor, shared epitope, ESR and CRP, and a baseline radiograph have significantly greater capacity to predict longterm radiographic progression compared to a patient questionnaire[18]. However, patient questionnaires have far greater capacity than a radiograph, joint count, or laboratory test to predict other severe patient outcomes of RA, including work disability, costs, premature death, and even joint replacement surgery[19].

The continuous PRO-Index C method removes arbitrary 20%, 50%, or 70% cutoff points seen with ACR response criteria and the categorical "majority" PRO-Index M. A continuous index usually provides greater statistical power than a categorical index[20]. Further, a continuous index has the capacity to recognize worsening, expressed as a possible negative change[9], in addition to improvement, as seen in Figure 3. For example, if one patient has a 25% improvement and a second patient a 25% deterioration, the net result of a continuous index would be 0. By contrast, according to ACR20 categorical criteria, the net result with 25% improvement and 25% deterioration would be expressed as a "20% improvement in 50% of patients." Most reports of RA clinical trials indicated that 20%–30% of patients met ACR20 response criteria with comparator or placebo treatment. Such results may be in part an artifact of the structure of ACR response criteria, rather than an actual direct benefit or a "placebo effect"[21-23].

Some limitations are seen in these studies. First, the total number of clinical trials that have been analyzed according to PRO indices remains small. Nonetheless, our report increases the number of clinical trials that have been analyzed according to PRO indices from 1 to 5. It would appear desirable that clinical trials involving other agents be analyzed to further characterize the possible value of PRO indices.

A second limitation is that all analyses are post-hoc. It is unlikely that prospective assessment in a clinical trial would be performed according to a PRO Index. Indeed, we do not suggest that any clinical trial be conducted only according to a PRO Index, as joint counts have greater specificity for change in RA clinical status than PRO measures[24]. Nonetheless, PRO Index data could be reported in addition to ACR20, 50, 70, ACR-N and/or DAS28 criteria, to provide prospective data concerning performance of such indices to distinguish active from control treatment compared to standard indices.

Perhaps the most important advantage of PRO indices is that quantitative and informative data can be acquired easily in standard clinical care, without joint counts or laboratory tests[25]. Most rheumatologists do not perform quantitative joint counts in most patients at most visits[7], and, therefore, do not have available quantitative documentation of clinical improvement unless with ESR or CRP, or patient questionnaire data. Even if patient questionnaire measures have somewhat less sensitivity than information in a full ACR Core Data Set or DAS28, their pragmatic utility might be desirable compared to the current situation in which most clinical rheumatology care is conducted without quantitative information[26]. The data reported here, and previous reports[9,10], further support that patient questionnaire data alone can also be used effectively to assess changes in clinical status in patients with RA in standard clinical care.

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. 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 1994;21 Suppl 41:86-89.
3. van der Heijde DMFM, van't Hof M, van Riel PLCM, van de Putte LBA. Development of a disease activity score based on judgment in clinical practice by rheumatologists. J Rheumatol 1993;20:579-81.
4. 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.
5. Wolfe F, Pincus T, Thompson AK, Doyle J. The assessment of rheumatoid arthritis and the acceptability of self-report questionnaires in clinical practice. Arthritis Care Res 2003;49:59-63.
6. Pincus T, Segurado OG. Most visits of most patients with rheumatoid arthritis to most rheumatologists do not include a formal quantitative joint count. Ann Rheum Dis 2006;65:820-2.
7. Wolfe F, Michaud K. The clinical and research significance of the erythrocyte sedimentation rate. J Rheumatol 1994;21:1227-37.
8. Siegel JN, Zhen BG. Use of the American College of Rheumatology N (ACR-N) index of improvement in rheumatoid arthritis: argument in favor. Arthritis Rheum 2005;52:1637-41.
9. Pincus T, Amara I, Koch GG. Continuous indices of Core Data Set measures in rheumatoid arthritis clinical trials: lower responses to placebo than seen with categorical responses with the American College of Rheumatology 20% criteria. Arthritis Rheum 2005;52:1031-6.
11. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies and clinical trials: the patient activity scale (PAS/PAS-II). J Rheumatol 2005;32:2410-5.
12. Weinblatt ME, Keystone EC, Furst DE, et al. Adalimumab, a fully human anti-tumor necrosis factor alpha monoclonal antibody, for the treatment of rheumatoid arthritis in patients taking concomitant methotrexate: the ARMADA trial. Arthritis Rheum 2003;48:35-45.
13. Keystone EC, Kavanaugh AF, Sharp JT, et al. Radiographic, clinical, and functional outcomes of treatment with adalimumab (a human anti-tumor necrosis factor monoclonal antibody) in patients with active rheumatoid arthritis receiving concomitant methotrexate therapy: a randomized, placebo-controlled, 52-week trial. Arthritis Rheum 2004;50:1400-11.
14. van de Putte LBA, Rau R, Breedveld FC, et al. Efficacy and safety of fully human anti-tumour necrosis factor a monoclonal antibody adalimumab (D2E7) in DMARD refractory patients with rheumatoid arthritis: a 12 week, phase II study. Ann Rheum Dis

2003;62:1168-77.
15. Furst DE, Schiff MH, Fleischmann RM, et al. Adalimumab, a fully human anti-tumor necrosis factor-a monoclonal antibody, and concomitant standard antirheumatic therapy for the treatment of rheumatoid arthritis: results of STAR (safety trial of adalimumab in rheumatoid arthritis). J Rheumatol 2003;30:2563-71.
16. 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.
17. Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JJ. Self-report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic, and laboratory measures. Ann Intern Med 1989;110:259-66.
18. Olsen NJ, Callahan LF, Brooks RH, et al. Associations of HLA-DR4 with rheumatoid factor and radiographic severity in rheumatoid arthritis. Am J Med 1988;84:257-64.
19. Pincus T, Brooks RH, Callahan LF. Prediction of long-term mortality in patients with rheumatoid arthritis according to simple questionnaire and joint count measures. Ann Intern Med 1994;120:26-34.
20. Anderson JJ, Bolognese JA, Felson DT. Comparison of rheumatoid arthritis clinical trial outcome measures: a simulation study. Arthritis Rheum 2003;48:3031-8.
21. Spiro HM. Doctors, patients, and placebos. New Haven: Yale University Press; 1986.
22. Gotzsche PC, Hansen M, Stoltenberg M, et al. Randomized, placebo controlled trial of withdrawal of slow-acting antirheumatic drugs and of observer bias in rheumatoid arthritis. Scand J Rheumatol 1996;25:194-9.
23. Kaptchuk TJ. Powerful placebo: the dark side of the randomised controlled trial. The Lancet 1998;351:1722-5.
24. Pincus T. The DAS is the most specific measure, but a patient questionnaire is the most informative measure to assess rheumatoid arthritis. J Rheumatol 2006;33:834-7.
25. Pincus T, Wolfe F. An infrastructure of patient questionnaires at each rheumatology visit: Improving efficiency and documenting care. J Rheumatol 2000;27:2727-30.
26. Pincus T, Wolfe F. Patient questionnaires for clinical research and improved standard patient care: is it better to have 80% of the information in 100% of patients or 100% of the information in 5% of patients? J Rheumatol 2005;32:575-7.