The American College of Rheumatology (ACR) Core Data Set and derivative "patient only" indices to assess rheumatoid arthritis

T. Pincus

Please address correspondence to: Theodore Pincus, MD, Division of Rheumatology and Immunology, Vanderbilt University School of Medicine, 203 Oxford House, Box 5, Nashville, TN 37232-4500, USA.
E-mail: t.pincus@vanderbilt.edu

Supported in part by grants from the Arthritis Foundation, the Jack C. Massey Foundation, and Abbott Immunology.

Clin Exp Rheumatol 2005; 23 (Suppl. 39): S109-S113.

Key words: Rheumatoid arthritis, "patient only" index, ACR 20, ACR 50, ACR 70.

ABSTRACT

Pooled indices of several measures have been developed to assess and monitor patients with rheumatoid arthritis in clinical trials and clinical care, as no single measure can serve as a "gold standard" in all individual patients. Early indices of disease activity include the Steinbrocker "therapeutic scorecard in rheumatoid arthritis," the Lansbury Index, and Paulus criteria. The most widely used indices at this time are the American College of Rheumatology (ACR) Core Data Set and disease activity score (DAS). A simplified disease activity index (SDAI) and clinical disease activity index (CDAI) are derived from the DAS. The ACR Core Data Set includes 7 measures – swollen joint count, tender joint count, patient assessment of global status, an acute phase reactant [erythrocyte sedimentation rate (ESR) or C-reactive protein (CRP)], health professional assessment of global status, physical function, and pain; the first four of these measures are included on the DAS. Improvement criteria for the ACR Core Data Set are based on improvement of at least 20% in both tender and swollen joint counts, and three of the five additional measures (ACR 20), and corresponding "ACR 50," and "ACR 70." A pooled index which includes only the three patient self-report questionnaire measures from the Core Data Set, physical function, pain, and patient assessment of global status performs as well as ACR 20 or DAS to discriminate between efficacy of active versus placebo treatment in a clinical trial.

Introduction

As discussed in the first chapter of this supplement (1), no single measure, such as blood pressure or serum cholesterol, can be applied as a "gold standard" to assess all individual patients with rheumatoid arthritis (RA) in clinical trials, clinical research and clinical care. Measures that are abnormal in the majority of patients may be normal in many patients. For example, the erythrocyte sedimentation rate (ESR) is normal in 40% of patients with RA(2), and some patients with RAmay report no pain.

The absence of a single "gold standard" measure has led to the development of pooled indices of several measures, which may be informative to assess responses in all patients. Goldsmith and colleagues documented in hypothetical clinical trials of 10 patients with RA that no single measure among active joints, ESR, morning stiffness, grip strength, and change score could discriminate between active or placebo treatment, but that a pooled index of these measures yielded significant results (3). Therefore, a pooled index was a more powerful approach to the assessment of clinical outcomes than these individual measures. In this chapter, indices to assess RA are reviewed briefly (Table I), not including the DAS (4) [see (5)] and SDAI (6) [see (7), which are reviewed elsewhere in this Supplement.

Historical indices of RAdisease activity

Early efforts to develop indices for the assessment of RA began almost 50 years ago, with "a therapeutic scorecard" for rheumatoid arthritis (8) (Table I). This index included joint swelling, joint limited motion, joint tenderness, the erythrocyte sedimentation rate (ESR), hemoglobin, weight, pain, global well-being and functional status. In retrospect, the therapeutic scorecard includes most of the measures regarded to be of importance in the assessment of RAto this day. The scorecard incorporates the American College of Rheumatology (ACR) Core Data Set, plus

Table I. Indexes of various measures used to analyze rheumatoid arthritis.

Steinbrocker Therapeutic Scorecard " (8):
1. Joint Swelling
2. Joint Motion
3. Joint Tenderness
4. Erythrocyte Sedimentation Rate
5. Hemoglobin
6. Weight
7. Pain
8. Well-being
9. Functional Capacity

Lansbury Systemic Manifestations of Rheumatoid Activity (11):
1. Erythrocyte Sedimentation Rate
2. Pain on Motion
3. Muscle Weakness
4. Morning Stiffness
5. Fatigability
6. Anemia
7. Pain at Rest
8. Fever

Paulus Criteria - requires 20% improvement in 4 of 6 measures (13):
1. Morning Stiffness
2. Westergren Erythrocyte Sedimentation Rate
3. Joint Pain/Tenderness Score
4. Joint Swelling Score
5. Five Point Patient Global
6. Five Point Physician Overall Assessment of Current Disease Severity

ACR Core Data Set (14-16):
1. Swollen Joint Count
2. Tender Joint Count
3. Physician Global Status
4. Acute Phase Reactant – ESR or CRP
5. Physical function
6. Pain
7. Patient Global
8. Radiograph, if study includes more than 1 year

Disease Activity Score (DAS) (4):
1. Swollen Joint Count
2. Tender Joint Count
3. Acute Phase Reactant – ESR or CRP
4. Patient Global Assessment

Simplified disease activity index (SDAI) (6)
1. Swollen Joint Count
2. Tender Joint Count
3. Acute Phase Reactant – ESR or CRP
4. Patient Global Assessment
5. Physician Global Assessment

Patient only index (23)
1. functional disability
2. pain
3. global status.

weight, hemoglobin and joint limited motion. However, it did not become widely used, perhaps in part due to the emphasis that came to be placed on laboratory science after the discovery of rheumatoid factor in 1940 (9) and 1948 (10).

The Lansbury Index (11) (Table I) includes ESR, joint pain on motion, muscle weakness, morning stiffness, fatigue, anemia, pain at rest and fever, with the joint indices being weighted in proportion to their surface area. While a weighted joint count appears to be a rational approach, weighted indices have not been found to be superior to standard indices (12). This method did not gain widespread use.

Paulus *et al*. (13) (Table I) developed an index consisting of morning stiffness, ESR, joint pain/tenderness score, joint swelling score and patient and physician global overall assessments on a 1-5 point scale. The "Paulus criteria" of a 20% improvement in 4 of the 6 of these measures were found to discriminate between the relative efficacies of DMARDs and placebo in clinical trials (13).

The ACR Core Data Set

These early indices (11-13) provided advances for rheumatology. However, up until the early 1990s, RAclinical trials and clinical research were characterized by many different outcome measures, ranging from grip strength to joint counts to laboratory measures. The situation suggested the need for a standard core data set of measures designed for use in RA clinical trials, and possibly in standard clinical care.

A committee was established by the American College of Rheumatology (ACR) to design a Core Data Set of measures for RA clinical trials (14). Initially, various measures from the literature were analyzed for construct, face, content, criterion, and discriminant validity (14). Sensitive candidate measures identified included the tender joint count, ESR, swollen joint count, physician global assessment, platelet count, grip strength, patient global assessment, pain, morning stiffness, hemoglobin, functional class, PIP circumference, walk time, quality of well-being, and digital joint size.

These results were compiled into a preliminary Core Data Set, which was presented at the first Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) meeting held in Maastricht, The Netherlands, to the international community of rheumatologists. Small groups reviewed paper cases to designate whether patients with changes in various candidate measures had better, the same, or a worsened clinical status from baseline to end point. The ACR committee reconvened after the OMERACT conference to reach a further consensus, which became the ACR Core Data Set (14).

The ACR Core Data Set (14-16) (Table I) includes 7 measures: 3 by an assessor swollen joint count, tender joint count

and physician assessment of global status, 3 by patient self-report physical function, pain and global status on patient questionnaire, and one acute phase reactant – ESR or C-reactive protein CRP). A radiograph is included for studies of one year or longer.

Further analyses led to the ACR preliminary definition of improvement (17) as improvement of at least 20% in both the tender and swollen joint counts, as well as 3 of the 5 additional measures; this came to be known as the "ACR 20". Higher thresholds for improvement such as "ACR 50," and "ACR 70" have also been described (18). The ACR 20 response was found to distinguish between active treatment with disease modifying anti-rheumatic drugs (DMARDs) versus placebo treatment more effectively than the ACR 50 and ACR 70 responses (17, 18).

Assessment of new therapies in RAhas been substantially improved through the ACR Core Data Set and the 20% 50% and 70% improvement criteria, as well as the disease activity score (DAS) (4) (see 5). The same measures are reported in all contemporary clinical trials, leading to a higher capacity of rheumatologists and regulatory agencies to interpret results. The convenience of reporting an underlying continuous index in several categories has proven satisfactory to recognize significant differences in the efficacy of active versus placebo treatments in clinical trials of new agents for RA, most notably biological agents. Results are also computed according to the DAS.

The ACR-N Index

The ACR-N Index (19) is derived from the ACR Core Data Set. Differences between baseline and endpoint are calculated for each of the 7 measures. However, rather than ascertaining whether a patient has a 20% (or 50% or 70%) improvement in the number of swollen joints, tender joints, and 3 of the other 5 measures, ACR-N reports a continuous variable, as the minimum of percent change in swollen joints, tender joints, and median of the other 5 Core Data Set measures. One advantage of ACR-N is that as a continuous variable, data may be reported at any level, rather than at arbitrary 20%, 50%, and 70% levels. A second advantage is that a continuous index can include recognition of possible deterioration in addition to improvement, which is not included in the ACR improvement criteria.

"Patient only" indices derived from the ACR Core Data Set

The relative efficiencies of the 7 ACR Core Data Set measures to detect differences between active and placebo responses in a randomized controlled clinical trial of leflunomide versus methotrexate versus placebo in patients with active RA was analyzed (20, 21). Two measures on a patient questionnaire, patient physical function and patient assessment of global status, had substantially higher relative efficiencies compared to the tender joint count. By contrast, the swollen joint count and erythrocyte sedimentation rate had lower relative efficiencies than the tender joint count. These observations may have been unexpected by most rheumatologists, who regard joint count measures as more valuable than patient questionnaire measures (22).

The high relative efficiency of the patient questionnaire measures suggested that a pooled index which included only the three patient-self-report questionnaire measures from the Core Data Set might perform as well as an ACR 20 or DAS in discriminating between the efficacy of active versus placebo treatments in a clinical trial. The advantage of a "patient only" index is that all data are provided by the patient, as most rheumatologists do not perform quantitative joint counts in most patients at most visits (22).

This possibility was tested using four analytic methods applied to a "patient only" index, based on scores for physical function, pain, and global status from data in the same clinical trial of leflunomide versus methotrexate versus placebo (20, 21). Four types of "patient only" indices were developed:

1. "Rescaled average": Based on rescaling each of the Core Data Set measures from 0 to 100, and computing an average value of 3 measures as the percent change from baseline to one year.
2. "Raw average": Computation of the composite average percent change from baseline for the 3 measures, without rescaling each measure.
3. "Categories": Based on percent changes for the included 3 measures as in the "raw average," but the percent changes are transformed to 5 ordered categorical variables: (-1) for more than 20% worsening from baseline; (0) for less than 20% worsening or improvement; (+1) for at least 20% but less than 50% improvement; (+2) for at least 50% but less than 70% improvement; and (+3) for at least 70% improvement. This classification of averages into categories is similar in spirit to the ACR20.
4. "Majority": Favorable responses are defined as at least a 20% improvement in 2 of 3 components of the indices, in a manner similar to the ACR 20 (23).

Each of these indices yielded results similar to the ACR 20, with 20% improvement in a range of 52% to 74% for patients randomized to leflunomide, 45% to 72% for those who took methotrexate, and 18% to 43% for those who took placebo (23). Pairwise kappa statistics ranged from 0.57 to 0.80, indicating good agreement. Therefore, an index of "patient only" measures provided similar levels of greater efficacy for leflunomide or methotrexate versus placebo as the ACR 20 (as well as the DAS) (23).

A further variant of a "patient only" index has been to adapt the "majority" method as a continuous, rather than a categorical index (24). As noted for ACR-N, one advantage of this approach is that change can be expressed at any level rather than at arbitrary 20%, 50%, and 70% levels. A second advantage is that a continuous index can include the recognition of possible deterioration in addition to improvement, which is not included in the ACR improvement criteria.

For example, an RAclinical trial which indicates that placebo treatment results in half the patients showing 25% improvement in Core Data Set Measures (including the swollen and tender joint counts) and half showing 25% worsening would be reported as "50% of patients met ACR 20 response criteria," rather than as a net mean improvement

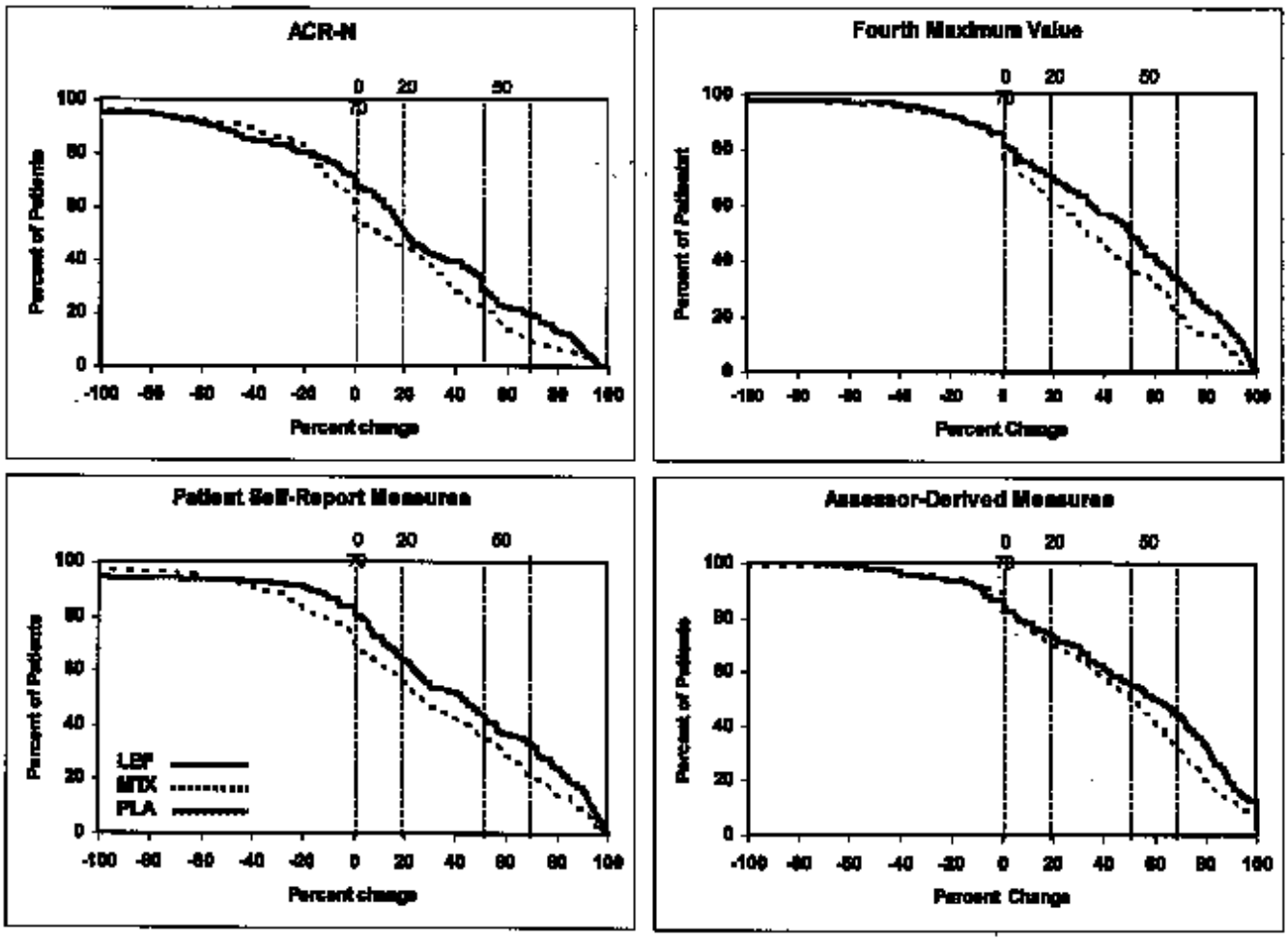

Fig. 1. Probability plots of the distributions of four continuous indices depicting the percent of patients who showed worsening or improvement after treatment with leflunomide, (dotted line), methotrexate (solid line) and placebo (dashed line). The indices are: ACR-N – lowest percent change of swollen joints, tender joints, and median of the other 5 measures; composite – median percent change of all 7 ACR Core Data Set measures; patient only" – median percent change of physical function, pain, global status; assessor only – median percent change of swollen joints, tender joints, global status.
These probability plots highlight the 0%, 20%, 50% and 70% levels depicted in Table II, but allow recognition of any level of responses, e.g., 50% responses are seen for 45% of patients with leflunomide, 36% with methotrexate and 13% with placebo for the "patient only" index. Note that positive results are seen with placebo for about only about 40% of patients according to ACR-N, about 50% for the composite or patient self-report indices, and about 60% for the assessor derived index.

of "0." This may overstate the efficacy of placebo for individual patients in the computation of group results; indeed, placebo treatment results in an apparent 15-30% responses according to ACR 20 and other indices in most RA clinical trials (13,18).

Therefore, the same clinical trial was analyzed according to a "majority continuous" index (24). The percent change was computed for each of the 7 Core Data Set measures, with changes more negative than -100% recoded to -100% so as to avoid extreme outliers of worsening (performed for fewer than 5% of patients). Mean and median scores according the "patient only" continuous index of 3 self-report measures were 36% and 43% in patients randomized to leflunomide, 26% and 27% in patients treated with methotrexate, and 0.4% and 2% in patients treated with placebo. The median values indicate that at least half of the patients had at least 43% improvement in at least 2 of the patient-derived measures for leflunomide, compared to 27% for methotrexate and 2% for placebo (24).

All differences between leflunomide versus placebo and methotrexate versus placebo were statistically significant ($p < 0.001$). Differences in patients meeting 20%, 50% and 70% improvement with active versus placebo treatment were higher for the "patient only" index than for ACR 20, 50 and 70 (24). Probability plots indicated that negative results indicative of clinical worsening were seen with placebo for about 50% of patients according to the patient self-report index, in contrast to 26% of placebo-treated patients with ACR 20 responses (Fig. 1) (24). Similar results were seen for ACR-N, "all Core Data Set" and "assessor only" indices (Fig. 1) (24).

These results raise consideration that the results of drug therapy in RAin clinical trials, observational studies, and routine clinical care might be assessed effectively using solely patient questionnaire data. Data from patient questionnaires are correlated significantly with data from traditional joint counts, radiographs, and laboratory tests (25) and are more explanatory of other clinical information than any other data in

RA (25). Patient questionnaire data are as effective as any available data in RA for the documentation of declines in functional status (26, 27), and to predict work disability (28), costs (29), and premature mortality (30).

The three Core Data Set measures are found on one page on the health assessment questionnaire (HAQ) (31), clinical health assessment questionnaire (CLIN-HAQ) (32), and one side of one page on the multi-dimensional health assessment questionnaire (MDHAQ) (33), which can be completed by a patient in 5-10 minutes or less in a waiting room (34). The MDHAQ can be scored in less than 20 seconds. Introduction of a simple patient questionnaire into clinical trials and the infrastructure of routine rheumatology care (35) would allow assessment based on a simple index in order to characterize quantitatively the status and responses to any therapy of patients with RA.

References

1. PINCUS T, SOKKA T: Complexities in assessment of patients with rheumatic diseases in clinical trials and clinical care. *Clin Exp Rheumatol* (in press).
2. WOLFE F, MICHAUD K: The clinical and research significance of the erythrocyte sedimentation rate. *J Rheumatol* 1994;21:1227-37.
3. GOLDSMITH CH, SMYTHE HA, HELEWA A: Interpretation and power of pooled index. *J Rheumatol* 1993; 20: 575-8.
4. 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.
5. FRANSEN J, VAN RIEL PLCM: The disease activity score and the EULAR response criteria. *Clin Exp Rheumatol*. In press.
6. SMOLEN JS, BREEDVELD FC, SCHIFF MH *et al*.: A simplified disease activity index for rheumatoid arthritis for use in clinical practice. *Rheumatology* (*Oxford*) 2003;42:244-57.
7. ALETAHA D, SMOLEN J: The simplified disease activity Index (SDAI) and the clinical disease activity index (CDAI): A review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol*. In press.
8. STEINBROCKER O, BLAZER A: A therapeutic score card for rheumatoid arthritis. *N Engl J Med* 1946; 14: 501-6.
9. WAALER E: On the occurrence of a factor in human serum activating the specific agglutination of sheep blood corpuscles. *Acta Path Microbiol Scand* 1940; 17: 172-8.
10. ROSE HM, RAGAN C, PEARCE E, LIPMAN MO: Differential agglutination of normal and sensitized sheep erythrocytes by sera of patients with rheumatoid arthritis. *Proc Soc Exp Biol Med* 1948; 68: 1-6.
11. LANSBURY J: A method for summation of the systemic indices of rheumatoid activity. *Am J Med Sci* 1956; 232: 300-10.
12. 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.
13. PAULUS HE, EGGER MJ, WARD JR, WILLIAMS HJ and the COOPERATIVE SYSTEMATIC STUDIES OF RHEUMATIC DISEASES GROUP: 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.
14. 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.
15. TUGWELL P, BOERS M: OMERACT Committee. Proceedings of the OMERACT Conferences on outcome measures in rheumatoid arthritis clinical trials, Maastrict, Netherlands. *J Rheumatol* 1993; 20: 527-91.
16. VAN RIEL PLCM: Provisional guidelines for measuring disease activity in clinical trials on rheumatoid arthritis (Editorial). *Br J Rheumatol* 1992; 31: 793-4.
17. 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.
18. FELSON DT, ANDERSON JI, LANGE MLM, WELLS G, LAVALLEY MP: Should improvement in rheumatoid arthritis clinical trials be defined as fifty percent or seventy percent improvement in core set measures, rather than twenty percent? *Arthritis Rheum* 1998; 41: 1564-70.
19. BATHON JM, MARTIN RW, FLEISCHMANN RM *et al*.: A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. *N Engl J Med* 2000; 343:1586-93.
20. STRAND V, COHEN S. SCHIFF M *et al*.: Treatment of active rheumatoid arthritis with leflunomide compared with placebo and methotrexate. *Arch Intern Med* 1999; 159: 2542-50.
21. TUGWELL P, WELLS G, STRAND V *et al*.: Clinical improvement as reflected in measures of function and health-related quality of life following treatment with leflunomide compared with methotrexate in patients with rheumatoid arthritis: Sensitivity and relative efficiency to detect a treatment effect in a twelve-month, placebo-controlled trial. Leflunomide Rheumatoid Arthritis Investigators Group. *Arthritis Rheum* 2000; 43: 506-14.
22. STRAND V, TUGWELL P, BOMBARDIER C *et al*.: Function and health-related quality of life: Results from a randomized controlled trial of leflunomide versus methotrexate or placebo in patients with active rheumatoid arthritis. *Arthritis Rheum* 1999; 42: 1870-8.
23. PINCUS T, STRAND V, KOCH G *et al*.: An index of the three core data set patient questionnaire measures distinguishes efficacy of active treatment from placebo as effectively as the American College of Rheumatology 20% response criteria (ACR20) or the disease activity score (DAS) in a rheumatoid arthritis clinical trial. *Arthritis Rheum* 2003; 48: 625-30.
24. 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.
25. 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.
26. PINCUS T, CALLAHAN LF, SALE WG, BROOKS AL, PAYNE LE, VAUGHN WK: Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984; 27: 864-72.
27. WOLFE F, CATHEYMA: The assessment and prediction of functional disability in rheumatoid arthritis. *J Rheumatol* 1991;18:1298-306.
28. WOLFE F, HAWLEY DJ: The longterm outcomes of rheumatoid arthritis: Work disability: A prospective 18 year study of 823 patients. *J Rheumatol* 1998; 25: 2108-17.
29. LUBECK DP, SPITZ PW, FRIES JF, WOLFE F, MITCHELL DM, ROTH SH: A multicenter study of annual health service utilization and costs in rheumatoid arthritis. *Arthritis Rheum* 1986; 29: 488-93.
30. 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.
31. FRIES JF, SPITZ P, KRAINES RG, HOLMAN HR: Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
32. HAWLEY DJ, WOLFE F: Sensitivity to change of the Health Assessment Questionnaire (HAQ) and other clinical and health status measures in rheumatoid arthritis: results of short term clinical trials and observational studies versus long term observational studies. *Arthritis Care Res* 1992; 5: 130-6.
33. PINCUS T, SWEARINGEN C, WOLFE F: Toward a multidimensional health assessment questionnaire (MDHAQ): Assessment of advanced activities of daily living and psychological status in the patient friendly health assessment questionnaire format. *Arthritis Rheum* 1999; 42: 2220-30.
34. PINCUS T: Documenting quality management in rheumatic disease: Are patient questionnaires the best (and only) method? *Arthritis Care Res* 1996; 9: 339-48.
35. PINCUS T, WOLFE F: An infrastructure of patient questionnaires at each rheumatology visit: Improving efficiency and documenting care. *J Rheumatol* 2000; 27: 2727-30.