Editorial

Outcome Issues for Rheumatoid Arthritis: What Do We Have and What Do We Want?

In this issue of *The Journal*, El Miedany, *et al*[1] have described a study examining the outcome of patients treated with anti-tumor necrosis factor-α therapy in terms of the ability of the Disease Activity Score 28 (DAS 28) to identify patients who have and who have not responded to the therapy. They examine various permutations of the DAS 28[2] in order to identify those aspects of the scoring system that were problematic. The gold standard was the ability to discriminate responders from nonresponders, and with this approach they were able to show that C-reactive protein (CRP) was a more accurate measure of disease activity than erythrocyte sedimentation rate, and that patient global assessment was inferior to physician global assessment. This is, in a sense, an internal examination of the validity of the components of the DAS 28 and doesn't address the larger question of what are and are not the characteristics of outcome measures that would be maximally useful to physicians and patients.

Starting from the premise of what do we want, it is self-evident that we need a scale that adequately discriminates between responders and nonresponders to a disease modifying drug[3]. A score should be reproducible, both by a single evaluator (test-retest reliability; that is, intrarater reliability) and across evaluators; that is, interrater reliability. It should be a continuous variable since the response of the individuals to medications is not an all-or-nothing phenomenon. In this sense the DAS 28 is a substantial improvement over the American College of Rheumatology response criteria ACR20, 50, or 70, although there have been many proposals to reconfigure the components of the ACR criteria as a continuous measure[4]. It should be relatively linear in the sense that a distribution of results falls on a continuum and should not be skewed, for example, if there are "floor" or "ceiling" effects. The measure should have equal ability to discriminate in an interventional study such as a randomized controlled trial or in an effectiveness study such as an observational cohort of patients. The scoring system ought to correlate, and in a sense predict, an anatomic outcome such as erosions detected on radiography or magnetic resonance imaging[5]. The scoring system should have the usual criteria for validity such as face validity and convergent validity and so on. Last, we should be able to aggregate the measures over time and utilize techniques such as area under the curve to best describe the clinical course of patients[6].

With this preamble in mind, what do we actually want to know about patients with rheumatoid arthritis? Certainly inflammation of the joints is paramount, both as a manifestation of the disease and as a target for therapy. Given the difficulty in performing reliable joint examinations, the validity of joint counts has been extensively explored. A DAS 28 count joint examination is probably sufficient for patients with RA. The presence of inflammatory serum markers is unquestionably a facet of the disease that gives important clues to the pathogenesis and pathophysiology. In most studies, CRP performs better than sedimentation rate, but other acute phase reactants could be used. For the moment, the CRP is probably appropriate. Less clear are the global measures. Neither the patient nor the physician global is particularly reliable in either test/retest situations or interrelater reliability. One might consider substituting a patient based pain measurement such as the visual analog scale as an albeit qualitative measure of discomfort associated with the disease. Unfortunately, external features such as fibromyalgia or neuropathy often contaminate this. Nonetheless, pain is a major clinical characteristic that motivates therapeutic intervention and must be included in a measure. For practical reasons, extensive questionnaires such as the McGill Pain Questionnaire are not useful as a component of an outcome measure. To this end, a measure of disability should be included because it correlates with other consequences of the disease such as employability and economic and financial decline[7].

Last, an often ignored but important issue is the practicality of using the scale for clinical encounters of a routine

See Short-term outcome after anti-TNF-α therapy in RA: Do we need to revise our assessment criteria?, *page 490*

nature. Most of these scales have been used in clinical trials, but a practical summary score of disease activity would be immensely useful in the clinic to quantify a patient's disease activity and response to therapy longitudinally.

The available measures, predominately the ACR20, 50, and 70 and the DAS 28, are useful but limited outcome measures for RA. They were designed primarily to evaluate new therapeutic modalities in a trial scenario and none has proved very useful in a day to day clinical setting. All the measures are given at one point in time and judged in a categorical fashion against pretreatment values. In addition, the ACR20, 50, and 70 are dichotomous scales, which tend to reduce the power of the continuous measures used as components for them. It might, however, be reasonable to retain a dichotomous (categorical) outcome of remission (no tender, swollen joint, AM stiffness, pain, fatigue, or elevated acute phase reactants) or clinical remission (no tender swollen joints or acute phase reactants)[8]. Both these designations have been proposed as ACR criteria with a rough correlation with DAS score of between 2.32 and 2.60[9].

Scale construction is an intricate and time consuming activity. The major judicial body that examines these outcome measures is OMERACT, which meets on a yearly basis to discuss potential revision of the available scales[10].

In summary, we are part way toward developing a truly useful outcome measure of RA. I believe that in the future patient and physician global assessment will be discarded. A pain measurement and the functional disability component of the ACR scoring system will be retained, but modified, so that the floor and ceiling effects of the Health Assessment Questionnaire are eliminated for less severely disabled individuals. There seems to be no substitute on the horizon for an adequate physical examination, which relegates assessment to experienced physicians and nurse practitioners. In addition, CRP test, which is our best choice among acute phase reactants, is usually not available on a same-day basis, so the score cannot be calculated "in real time." Eventually, we will require a "real-time," reliable, and valid instrument that can be used in ongoing longitudinal care of patients with rheumatoid arthritis: a useful instrument in the clinic and in studies that will permit us to extrapolate directly from studies to our patients. It may even clarify what clinical characteristics are most likely to be addressed by which therapeutic modality.

DANIEL A. ALBERT,
5 Maloney Building, Suite 504,
3400 Spruce Street,
Philadelphia, Pennsylvania 19104-4283, USA

Address reprint requests to Dr. Albert.
E-mail: albertd@mail.med.upenn.edu

REFERENCES

1. El Miedany Y, Youssef SS, El Gaafary M. Short-term outcome after anti-TNF-alpha therapy in RA: Do we need to revise our assessment criteria? J Rheumatol 2006;33:490-6.
2. 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.
3. Buchbinder R, Bombardier C, Yeung M, Tugwell P. Which outcome measures should be used in rheumatoid arthritis clinical trials? Clinical and quality-of-life measures' responsiveness to treatment in a randomized controlled trial. Arthritis Rheum 1995;38:1568-80.
4. 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.
5. van der Heijde DM, van Riel PL, van Leeuwen MA, van 't Hof MA, van Rijswijk MH, van de Putte LB. Prognostic factors for radiographic damage and physical disability in early rheumatoid arthritis. A prospective follow-up study of 147 patients. Br J Rheumatol 1992;31:519-25.
6. Albert DA, Huang G, Dubrow G, Brensinger CM, Berlin JA, Williams HJ. Criteria for improvement in rheumatoid arthritis: alternatives to the American College of Rheumatology 20. J Rheumatol 2004;31:856-66.
7. Kobelt G, Eberhardt K, Jonsson L, Jonsson B. Economic consequences of the progression of rheumatoid arthritis in Sweden. Arthritis Rheum 1999;42:347-56.
8. van Riel PL, Fransen J. To be in remission or not: is that the question? Ann Rheum Dis 2005;64:1389-90.
9. Makinen H, Kautiainen H, Hannonen P, Sokka T. Is DAS28 an appropriate tool to assess remission in rheumatoid arthritis? Ann Rheum Dis 2005;64:1410-3.
10. Felson DT, Anderson JJ, Boers M, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis Rheum 1993;36:729-40.