Editorial

Longterm Treatment Benefits Are Best Reflected by Patient Reported Outcomes

Since 1998, the American College of Rheumatology (ACR) criteria (ACR 20/50/70% responses) and improvements in the Disease Activity Score (DAS) have been utilized to demonstrate efficacy of 7 new disease modifying antirheumatic drugs (DMARD). "Improvement in physical function and health related quality of life [HRQOL]" are established labeling claims, in response to US Food and Drug Administration requirements for "durability of response" over 24 months of treatment. The Health Assessment Questionnaire (HAQ) has become the primary measure of physical function in RA, accompanied by use of generic measures such as the Medical Outcomes Study Short-Form 36 (SF-36), EuroQOL (EQ-5D), and Health Utilities Index-3 (HUI3) to assess HRQOL in RA[1]. Patient reported outcomes have been shown to best differentiate active from placebo therapy[2-4].

In this issue of *The Journal*, Mittendorf and colleagues report treatment associated improvements in HRQOL and fatigue in 505 patients with RA who successfully completed several randomized dose-finding protocols and received longterm treatment with adalimumab for a mean of 1.6 years[5]. Mean disease duration in this population was 12.4 years, and patients had failed a mean of 3.7 prior DMARD; 47% were retired and only 30% were employed. The largest cohort in this open-label continuation study had completed the placebo controlled 6 month monotherapy DE011 trial. Those receiving 40 mg adalimumab (n = 96–99) or placebo (n = 91–93) every other week provide a comparison of initial responses to active therapy.

Mean baseline SF-36 scores in DE011 were low, indicating that patients with longer disease duration report more decrements in role-physical, physical functioning, bodily pain, and vitality. Treatment-associated improvements in SF-36 scores confirmed rapid and sustained improvements in HRQOL over 26 weeks, from 25.4 to 34.4 and 43.7 to 50.0 in physical component (PCS) and mental component scores (MCS), respectively, which were consistent with baseline values in the entire DE033 cohort. After longterm treatment with this tumor necrosis factor (TNF) inhibitor, mean area-under-the-curve of SF-36 scores demonstrated that these clinically meaningful improvements in HRQOL were sustained. HUI3 scores were low at baseline, indicating impairment due to both RA and comorbid conditions. The rapid and sustained improvements with adalimumab treatment add additional support of the sensitivity of the HUI3[6].

This is not entirely new information. The US301 randomized controlled trial (RCT) comparing leflunomide or methotrexate (MTX) with placebo first demonstrated that SF-36, a well known and validated generic measure of HRQOL, was sensitive to change in RA, and reflected longterm improvement over 2 years of treatment with either DMARD[2,3,7]. The randomized controlled ATTRACT trial demonstrated sustained improvements in physical function and HRQOL, over longterm treatment with infliximab plus MTX, in patients with longstanding disease duration who had failed multiple DMARD, including MTX[8,9]. And in TEMPO, mean improvements in HAQ, global disease activity, general health, and EQ-5D were highly correlated over 1 and 2 years of treatment[10].

It is important to note that longterm treatment with adalimumab similarly offers sustained benefit, again in an RA population with long disease duration having failed multiple DMARD. Although specific "n" are not available in the figures or tables to indicate the number of patients at each timepoint, about 73% completed questionnaires after approximately 3 years of treatment. The authors demonstrated high correlations between PCS as assessed by SF-36, physical function by HAQ, and fatigue by Functional Assessment of Chronic Illness Therapy-Fatigue scale (FACIT) and by HUI3 in this population at baseline. It would have been useful to show that treatment-associated improvements in these instruments were also correlated, and that the relative sensitivity and specificity of treatment-association changes with each instrument.

See Improvement and longterm maintenance of QOL during treatment with adalimumab in severe RA, *page 2343*

What lessons are to be learned from this article and what questions does it raise?

1. Another demonstration that active treatment with a TNF inhibitor, adalimumab, results in rapid improvements in physical function, HRQOL, and fatigue within 6 months that are sustained for a mean of 1.6 years in those continuing therapy.
2. That SF-36, a generic measure, demonstrates that HRQOL is significantly affected in patients with longstanding RA, across all domains, but particularly role-physical, physical functioning, bodily pain, and vitality. Active treatment results in statistically significant and clinically meaningful improvements in domain and PCS and MCS scores — yet after treatment all but mental health and MCS values still reflect large decrements from age and gender-matched US normative data.
3. That each of these instruments resulted in statistically significant and clinically meaningful improvements, defined as mean improvements within the group that met or exceeded "minimally clinically important differences" (MCID)[11,12].
4. Reported improvements in HRQOL were consistently reflected by FACIT and HUI3. The effect of active disease causing fatigue has been emphasized by the patient outcomes group of OMERACT, and several validated measures of fatigue are available for use in RA[13,14]. Do changes in FACIT correlate with improvement in the SF-36 vitality domain? Is additional information offered by use of both instruments? Similarly, SF-36 may be utilized to calculate utilities by SF-6D and may be shown to be as sensitive as HUI3[6]. Would these results correlate with reported improvements using HUI3?
5. As expected, each of these instruments is highly correlated at baseline, and with DAS28 scores. Does each offer valuable information? Or is an "economy of effort" possible so that fewer instruments may be used? The authors raise these questions at the conclusion of their report and it is hoped that this database will be used to help address these issues.

Finally, if we want to "distill" results of randomized controlled trials to best tell how patients report improvements that are "important to them," which instruments are best? Do they tell patients and practitioners which questions to ask? How may we translate trial data into by-patient results that can best inform clinical practice? How are "clinically meaningful" improvements best defined?

The concept of MCID has helped, particularly in allowing us to know the percentage of patients with improvements that meet or exceed this "minimum" value. Changes of 5–10 points in SF-36 domain scores and 2.5–5.0 points in PCS and MCS summary scores are considered to represent MCID, based on correlations with patient global assessments of disease activity and/or the Guyatt "feeling thermometer" in multiple rheumatologic diseases[15-18]. Similarly, the MCID can indicate those who attain a predefined "goal" such as HAQ scores of "0", or "0.5", representing population normative data[19], and/or SF-36 domain or PCS, which approach age and gender-matched US norms. Further, "really important differences" have been proposed to offer a more clinically meaningful definition of improvement[20]. Median changes within treatment groups can also be utilized, indicating that a majority of patients have achieved improvements greater than the MCID. Or, improvements ≥ 0.5 standard deviations of the mean change, considered to represent "minimally important differences," a statistical definition of improvement that is not specifically anchored to patient reported outcomes[21].

The article by Mittendorf and colleagues[5] again points out that SF-36, a generic measure of HRQOL, reflects the significant effect of RA on different aspects of patients' lives and functioning within their physical, emotional, and social environment — and sustained benefits resulting from longterm treatment. This raises an important question: Can we adequately assess treatment-associated improvements in disease activity in RCT by examining, not just physical function and/or fatigue, but *without* their influence across multiple domains of HRQOL? The value of estimating utilities by SF-6D and HUI3 is increasingly important, as are efforts to better understand work productivity in this disease population[22]. The richness of this dataset (and others available), which includes multiple instruments, will allow assessment of the relative sensitivity and specificity of treatment-associated changes with each instrument. This will facilitate choice of the most meaningful and pragmatically utilized patient reported outcomes for day to day use in clinical practice.

C. VIBEKE STRAND, MD,
Clinical Professor, Adjunct,
Division of Immunology/Rheumatology,
Stanford University,
Palo Alto, California;
BRUCE CRAWFORD, MA, MPH,
Mapi Values,
Boston, Massachusetts, USA

Address reprint requests to Dr. Strand. E-mail: vstrand@stanford.edu

REFERENCES

1. Marra CA, Woolcott JC, Kopec JA, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D) and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med 2005;60:1571-82.
2. Tugwell P, Wells G, Strand V, et al. Clinical improvement as reflected in measures of function and health related quality of life: sensitivity and relative efficiency to detect a treatment effect in a 12 month placebo controlled trial comparing leflunomide with methotrexate. Arthritis Rheum 2000;43:506-14.
3. Strand V, Cohen S, Crawford B, Smolen JS, Scott DL. Patient-reported outcomes better discriminate active treatment from placebo in randomized controlled trials in rheumatoid arthritis. Rheumatology Oxford 2004;43:640-7.
4. Cohen S, Strand V, Aguilar D, Ofman JJ. Patient versus physician-

reported outcomes in rheumatoid arthritis patients treated with recombinant interleukin-1 receptor antagonist (anakinra) therapy. Rheumatology Oxford 2004;43:648-56.
5. Mittendorf T, Dietz B, Sterz R, Kupper H, Cifaldi MA, von der Schulenburg J-M. Improvement and longterm maintenance of quality of life during treatment with adalimumab in severe rheumatoid arthritis. J Rheumatol 2007;34:xxxx.
6. Marra CA, Rashidi AA, Guh D, et al. Are indirect utility measures reliable and responsive in rheumatoid arthritis patients? Qual Life Res 2005;14:1333-44.
7. Strand V, Scott DL, Emery P, et al. Physical function and health related quality of life: analysis of 2-year data from randomized, controlled studies of leflunomide, sulfasalazine, or methotrexate in patients with active rheumatoid arthritis. J Rheumatol 2005; 32:590-601.
8. Maini RN, Breedveld FC, Kalden JR, et al. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. Sustained improvement over two years in physical function, structural damage, and signs and symptoms among patients with rheumatoid arthritis treated with infliximab and methotrexate. Arthritis Rheum 2004;50:1051-65.
9. Strand V. Longer term benefits of treating rheumatoid arthritis: assessment of radiographic damage and physical function in clinical trials. Clin Exp Rheumatol 2004;22:S57-S64.
10. van der Heijde D, Klareskog L, Singh A, et al. Patient reported outcomes in a trial of combination therapy with etanercept and methotrexate for rheumatoid arthritis: the TEMPO trial. Ann Rheum Dis 2006;65:328-34.
11. Kosinski M, Zhao SZ, Dedhiya S, Osterhaus JT, Ware JE Jr. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. Arthritis Rheum 2000;43:1478–87.
12. Strand V, Bombardier C, Maetzel A, Scott D, Crawford B. (2001a). Use of minimum clinically important differences [MCID] in evaluating patient responses to treatment of RA [abstract]. Arthritis Rheum 2001;44 Suppl:S187.
13. Kirwan JR, Minnock P, Adebajo A, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007;34:1174-7.
14. Kirwan JR, Hewlett S. Patient perspective: reasons and methods for measuring fatigue in rheumatoid arthritis. J Rheumatol 2007;34:1171-3.
15. Zhao SZ, Feichtner JI, Tindall EA, et al. Evaluation of health related quality of life of rheumatoid arthritis patients treated with celecoxib. Arthritis Care Res 2000;13:112–21.
16. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol 2000;27:2635-41.
17. Strand V, Crawford B. Improvement in health-related quality of life in patients with systemic lupus erythematosus following sustained reductions in anti-dsDNA antibodies. Expert Rev Pharmacoeconomics Outcomes Res 2006;5:317-26.
18. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999;15:141-55.
19. Krishnan E, Sokka T, Hakkinen A, Hubert H, Hannonen P. Normative values for the Health Assessment Questionnaire disability index: benchmarking disability in the general population. Arthritis Rheum 2004;50:953–60.
20. Wolfe F, Michaud K, Strand V. Expanding the definition of clinical differences: from minimally clinically important differences to really important differences. Analyses in 8931 patients with rheumatoid arthritis. J Rheumatol 2005;32:583-9.
21. Angst F, Aeschlimann A, Stucki G. Smallest detectable and minimal clinically important differences of rehabilitation intervention with their implications for required sample sizes using WOMAC and SF-36 quality of life measurement. Ann Rheum Dis 2001;60:834-40.
22. Escorpizo R, Bombardier C, Boonen A, et al. Worker productivity outcome measures in arthritis. J Rheumatol 2007;34:1372-80.