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

VIBEKE STRAND, DAVID L. SCOTT, PAUL EMERY, JOACHIM R. KALDEN, JOSEF S. SMOLEN, and GRANT W. CANNON, for the Leflunomide Rheumatoid Arthritis Investigators Groups, PETER TUGWELL, and BRUCE CRAWFORD

ABSTRACT. Objective. To determine whether improvements in physical function and health related quality of life (HRQOL) are sustained over 2 years of blinded treatment with leflunomide (LEF), methotrexate (MTX), or sulfasalazine (SSZ) in patients with active rheumatoid arthritis (RA).

Methods. Three phase III randomized, controlled trials compared LEF, MTX, and SSZ in patients with active RA. Improvements in physical function were assessed by Health Assessment Questionnaire Disability Index (HAQ-DI) and Modified Health Assessment Questionnaire (MHAQ); monthly MHAQ and mean HAQ scores were used to calculate American College of Rheumatology responses; HAQ-DI was assessed at baseline and 6-month intervals. In US301, the Medical Outcomes Study 36-Item Short-Form questionnaire (SF-36) assessed treatment-associated changes in HRQOL at baseline and 6-month intervals.

Results. Mean and median improvements in HAQ-DI after 12 and 24 months of active treatment in all phase III protocols significantly exceeded –0.22 or a minimum clinically important difference (MCID). These improvements closely reflected positive changes in SF-36 that met or exceeded MCID in all domains with LEF and MTX treatment. Problem Elicitation Technique Top 5 scores reflected improvements in performance of physical activities most important to patients.

Conclusion. Improvements in physical function were sustained over 24 months of successful treatment with LEF, MTX, and SSZ, and reflected improvements in mental as well as physical domains of HRQOL. (J Rheumatol 2005;32:590–601)

Key Indexing Terms:
RHEUMATOID ARTHRITIS
HEALTH-RELATED QUALITY OF LIFE
LEFLUNOMIDE

From the Division of Immunology, Stanford University School of Medicine, Palo Alto, California, USA; Department of Rheumatology, Dulwich Hospital, London, UK; Department of Rheumatology and Rehabilitation, University of Leeds School of Medicine, Leeds, UK; Medizinische Klinik III Mit Poliklinik, Friedrich-Alexander-Universitat Erlangen-Nurnberg, Erlangen, Germany; Center for Rheumatic Diseases at Laing Hospital, Vienna, Austria; Veterans Affairs Medical Center and University of Utah, Salt Lake City, Utah, USA; Ottawa Hospital, Ottawa, Ontario, Canada; and Mapi Values, Boston, Massachusetts, USA.

These clinical trials were funded by Aventis Pharmaceuticals, Bridgewater, NJ, USA.

V. Strand, MD, Adjunct Clinical Professor, Division of Immunology, Stanford University School of Medicine; D. Scott, MD, Professor of Clinical Rheumatology, Department of Rheumatology, Dulwich Hospital; P. Emery, MD, Department of Rheumatology and Rehabilitation, University of Leeds School of Medicine; J. Kalden, MD, Medizinische Klinik III Mit Poliklinik, Friedrich-Alexander-Universitat Erlangen-Nurnberg; J. Smolen, MD, Center for Rheumatic Diseases at Laing Hospital; G. Cannon, MD, Veterans Affairs Medical Center and University of Utah; P. Tugwell, MD, Head, Department of Medicine, Ottawa Hospital; B. Crawford, MA, MPH, Director, Patient Reported Outcomes and Regulatory Consulting, Mapi Values.

Address reprint requests to B. Crawford, Mapi Values, 15 Court Square, Suite 620, Boston, MA 02108. E-mail: bruce.crawford@mapivaluesusa.com


Active rheumatoid arthritis (RA) leads to impairment in physical function, limiting activities and decreasing health related quality of life (HRQOL). It is easy to underestimate the impact of active RA on direct and indirect medical costs including the intangible effects on physical, emotional, and social functioning1-2. Patient-reported measures of physical function, pain, and global disease activity best discriminate active from placebo treatment in randomized controlled trials (RCT), reflect changes in disease severity over time, and best predict long-term morbidity and mortality3-9.

Several well validated and cross-culturally translated patient-reported questionnaires measure physical function in RA, including the Health Assessment Questionnaire (HAQ), modified HAQ (MHAQ), multidimensional HAQ (MDHAQ), the Arthritis Impact Measurement Scales (AIMS), and the Problem Elicitation Technique (PET), where patients rate physical activities most important to them, that are affected by their disease, and that they want improved by treatment10-12. To facilitate comparisons across other chronic diseases and pharmacoeconomic analyses,
generic HRQOL instruments have been incorporated in recent RCT in RA, commonly the Medical Outcomes Study (MOS) Short Form-36 (SF-36)\textsuperscript{13-19}.

As part of the clinical development program for leflunomide (LEF), a disease modifying antirheumatic drug (DMARD), patients enrolled in 3 phase III RCT continued blinded treatment, to demonstrate sustained benefits over 2 years. US301 was a 24-month trial offering patients alternative active treatment following documented lack of efficacy on or after 4 months of protocol participation. It was the first to define 3 co-primary endpoints of efficacy: signs and symptoms of active RA by American College of Rheumatology (ACR) responses at 12 months, radiographic progression at 12 months, and sustained improvement in physical function and HRQOL at 24 months\textsuperscript{20,21}. Similar analyses were conducted in the European protocols, where patients continued blinded treatment in extension protocols: MN303 (6 months) and MN305 (additional 12 months) after completion of MN301, and MN304 (12 months) after MN302\textsuperscript{22-25}. All protocols were designed in 1993; the European RCT were initiated in 1994 and US301 in 1995. As reported, significant improvements in physical function and HRQOL were evident at 6 and 12 months following LEF, methotrexate (MTX), and sulfasalazine (SSZ) administration, compared with placebo\textsuperscript{11,13,26}. Mean improvements in HAQ Disability Index (HAQ-DI) and ACR20/50 responses were consistent in patients receiving LEF (the only treatment group in all 3 RCT) across protocols, despite different baseline demographics, disease characteristics, and HAQ scores in each trial\textsuperscript{20,22,24}.

These analyses were prospectively defined to determine whether improvements in physical function and HRQOL after 6 and 12 months were sustained over 2 years of continued blinded therapy with LEF, MTX, or SSZ. In US301, anticipating low completion rates, data after 12 months' treatment with placebo were prospectively excluded. A secondary objective compared improvements in physical function and HRQOL between active therapies. All treatments remained blinded as LEF was approved in the US in 1998 and Europe in 1999; most patients completed 24 months' treatment, entering further extension studies before LEF was commercially available.

MATERIALS AND METHODS

Protocol designs and study populations. Detailed results from the phase III RCT have been published\textsuperscript{20-25}. Protocol US301 enrolled patients with active RA who had never received MTX (n = 190 LEF, 190 MTX, and 128 placebo), including 26 from Canada (n = 8 LEF, 8 MTX, and 10 placebo) who were not part of the initial publications\textsuperscript{21,27}. Protocols MN301/303/305 compared LEF or SSZ treatment (n = 133 each) with placebo (n = 92), in SSZ-naive patients in Europe, Australia, and South Africa\textsuperscript{22,23,25}. Protocol MN302/304 compared active treatment with LEF (n = 501) and MTX (n = 498) in MTX-naive patients in Europe and South Africa\textsuperscript{24}. All were powered to demonstrate efficacy by ACR response criteria. In all 3 RCT, LEF treatment included a 3-day loading dose of 100 mg/day followed by 20 mg daily doses. MTX was initiated at 7.5 mg/wk, increased to 15 mg/wk for Weeks 6–9, and 20 mg/wk in the second year of US301, and to 12.5–15 mg/wk in MN302/304, at the discretion of the investigator. Mean weekly MTX doses in US301 were 11.7 mg and 12.6 mg in Years 1 and 2, respectively; median doses 15 mg/wk. Mean weekly MTX doses in MN302/304 were 11.9 mg and 12.2 mg in Years 1 and 2; median doses 10 mg/wk. Folate administration was mandated in US301, and prescribed in 10% of patients in MN302/304. In MN301/303/305, SSZ was initiated at 500 mg and increased to a maximum of 2 g/day.

Analyses included all patients in US301 who completed 12 months’ treatment with at least one subsequent protocol visit (up to and including 24 months of therapy), missing observations in Year 2 were replaced using last observation carried forward (LOCF). The other “Year 2 cohorts” included all patients who entered extension protocols MN305 or MN304 with at least one protocol visit after study entry, e.g., in Year 2.

Measures of physical function and HRQOL. In US301, MHAQ was administered monthly as part of the ACR response criteria\textsuperscript{28}. At baseline and 6-month intervals and/or at treatment end, HAQ, PET, SF-36, MOS current health perceptions scale, and items addressing work productivity from the 1994 National Opinion Research Center Survey were completed. The MN protocols utilized HAQ on a monthly basis; mean total scores were calculated as components of the ACR response criteria. Across the 3 protocols, mean changes from baseline in the HAQ-DI were compared at baseline and at 6 (US301, MN301, MN302), 12 (US301, MN303, MN302), and 24 months (US301, MN305, MN304).

In US301, changes in HAQ-DI were correlated with the PET Top 5 scores and the SF-36 domain and summary scores. Of interest, the PET responses identified a broad range of physical activities as “most important,” reflecting the heterogeneity of RA populations\textsuperscript{13}. The SF-36 was selected for use in US301 because of its proven reliability in a large number of clinical conditions, although it had not yet been shown to be sensitive to treatment effects in RA RCT\textsuperscript{29}. It assesses 8 domains of HRQOL, scored from 0 (worst) to 100 (best), and physical and mental component summary scores (PCS and MCS, respectively) derived to have mean values of 50 and standard deviations of 10 in the general population\textsuperscript{30,31}. The Work Limitations Questionnaire included in the 1994 National Opinion Research Center Survey assessed work productivity. Patients ranked difficulties performing work-related activities on a scale ranging from 1 (none) to 6 (can’t do). Applying algorithms similar to the SF-36, scores range from 0 to 100, with higher scores indicating greater productivity\textsuperscript{32,33}.

Changes from baseline were compared with published values for minimum clinically important differences (MCID) identifying improvements perceptible to patients. Changes of -0.22 points (range -0.19 to -0.43) in HAQ-DI, -5.0 points in the PET Top 5, 5 to 10 points in domain, and 2.5 to 5 in PCS and MCS scores of the SF-36 are considered to reflect MCID\textsuperscript{16,34-41}.

Statistical methods. Primary analyses compared whether improvements in physical function and HRQOL in patients receiving active therapy with LEF, MTX, or SSZ at 12 months were sustained over 24 months of treatment in Year 2 cohort populations. Equivalence with active treatment over 12–24 months was recognized if 95% confidence intervals (CI) for differences between treatments included 0. Secondary analyses compared improvements in patients receiving LEF, MTX, and SSZ.

Before performing these analyses, a preliminary assessment calculated physical function and HRQOL scores for each treatment period, and tested...
important baseline variables between treatment groups. To control for inflated α risk (type I error due to multiple testing), initial analyses included a test of treatment-group effect on change scores at endpoint. The omnibus test was a multivariate analysis of variance (MANOVA) applied on differences of scores. Dependent variables were the vector of physical function/HRQOL differences at endpoint; independent variables were treatment, region, treatment × region, and other differences identified among study groups at baseline. These were performed as endpoint analyses in all patients entering Year 2 to fully assess effects of active treatment. The vector of differences in reported physical function and HRQOL at study exit, incorporated in the US301 omnibus test, consisted of all 8 domains of the SF-36, HAQ-DI, and MOS current health score. Only the HAQ-DI was utilized for protocols MN305 and MN304.

Comparisons of mean changes from baseline in LEF compared to MTX or SSZ treatment groups used a general linear model (GLM), including treatment, region, treatment × region, and any differences identified at baseline as factors. The 95% CI for differences between adjusted mean changes in LEF and MTX or SSZ treatment groups were calculated. Demographic and disease characteristics were tested by chi-square for categorical data and analysis of variance (ANOVA) for continuous data. Baseline SF-36 and HAQ scores were examined for potential floor and ceiling effects to ensure the instruments were sensitive to change at the extremes. Psychometric properties of the European HAQ were evaluated to ensure that the empirical structure of the HAQ corresponded to its hypothesized structure and was cross-culturally stable and appropriate for pooling. This psychometric evaluation consisted of a multi-trait attribute analysis and assessment of clinical validity.

In all 3 phase III protocols, data were shown to be non-normally distributed. Accordingly, the van Elteren extension to the Wilcoxon rank-sum test was used for continuous variables, and the Cochran-Mantel-Haenszel chi-square test for categorical variables, controlling for region in both tests.

**RESULTS**

Figure 1 presents the disposition of patients in each protocol and extension studies. Demographics and baseline disease characteristics for intention-to-treat (ITT) and Year 2 cohorts are presented in Table 1. In US301, of 190 randomized to receive LEF, 25 nonresponders entered blinded alternative therapy and received MTX; 35 of 190 switched from MTX to LEF; and 56 of 128 receiving placebo started LEF. Including alternative therapy, 65% to 72% of patients who were randomized to receive LEF and MTX, respectively, completed protocol treatment.

Although the Year 2 cohorts were enriched with responders, they were not composed exclusively of patients doing well: ACR20 responses at 12 months ranged from 60% to 77% across active treatment groups. Patients who chose not to enter extension protocols (MN303/5 and MN304) were approximately evenly divided between ACR responders and nonresponders (Figure 1). A high percentage of patients in the Year 2 cohorts completed 24 months of protocol treatment with all active therapies: LEF, 88% (MN305 and MN304); and SSZ, 78% (MN305); and SSZ, 78% (MN305). While there were notable differences between protocol populations with respect to baseline demographics and disease characteristics, including HAQ-DI values, no differences were evident between treatment groups within each protocol, or between the ITT and Year 2 cohorts (Table 1).
In the Year 2 cohort in US301, HAQ data were missing for only one patient receiving LEF, as well as SF-36 data for 5 with LEF and one with MTX. These missing data would not be expected to affect the results, and sensitivity analyses showed the data to be robust[^27]. In the MN305 and MN304 protocols, 15% to 19% and 8% to 9% of patients, respectively, were excluded because a validated Yugoslavian translation of the HAQ was not available when the RCT were initiated (n = 20), and 9 questionnaires were excluded due to inconsistent responses. These patients did not differ in demographic or baseline characteristics from the ITT or Year 2 cohorts[^25]. In all 3 RCT, baseline HAQ-DI scores did not indicate floor or ceiling effects. A potential floor effect was found only for the role-physical domain of the SF-36: 49% in US301 reported 0 scores at baseline, reflecting difficulties with work/other daily activities as a result of physical health. A potential ceiling effect was evident in the role-emotional domain: 46% of patients scored the highest possible value at baseline, implying few problems with work/other activities as a result of emotional health.

**US301 results.** In the Year 2 cohort, improvements from baseline in HAQ-DI evident at 6 months were sustained over 24 months of treatment (Table 2, Figure 2): –0.55 to –0.61 with LEF compared with –0.34 to –0.37, respectively, with MTX (p = 0.005 at 24 mo)[^21]. These are consistent with values reported in the ITT population at 12 months [–0.45 (n = 166) vs –0.26 (n = 169)][^13,20]. A majority of patients in both Year 2 cohorts (71% LEF, 59% MTX) reported improvements that met or exceeded MCID. Mean changes in PET Top 5 score from baseline to Months 12 and 24 in LEF and MTX treated patients were –9.5 and –9.1 vs –4.5 and –4.3, respectively, (p < 0.01 at 24 mo; Table 3) and numerically better than the ITT population at 12 months[^13,20]. Improvement in PET Top 5 scores met or exceeded MCID in 68% of patients receiving LEF.

At baseline, in US301, domain scores in SF-36 were
lower than age and sex matched US normative populations\textsuperscript{30}. Reported improvements in 3 of 8 domains of the SF-36 (bodily pain, vitality, and role-emotional) were statistically better ($p \leq 0.05$) with LEF than MTX; they were equivalent in other domains (Figure 3). Mean change scores reported by patients receiving LEF or MTX met or exceeded MCID by improvements in SF-36 PCS scores in 80\% LEF and 77\% MTX-treated patients. Mean changes in SF-36 PCS from baseline to 12 and 24 months in the US301 Year 2 cohort with LEF and MTX, 35\% and 28\%, respectively, were not statistically different (Table 3). Baseline SF-36 MCS scores approached US norms of 50 in both active
treatment groups (48.5 and 49.8, respectively); mean treatment-associated changes were small, 4.7 (10%) and 2.7 (5%), respectively. In the ITT and Year 2 cohorts, improvements in current health perceptions were equivalent with LEF and MTX treatment, as were positive changes in work productivity (Table 3).

**MN301/303/305 results.** Mean changes from baseline in the HAQ-DI at 6 and from 12 to 24 months (~0.73 in the LEF and ~0.56 in the SSZ treatment groups at 24 mo) were not statistically different; mean percentage improvements from baseline to Month 24 were numerically greater with LEF compared with SSZ treatment (46% vs 37%; Figure 2, Table 2). Improvements in physical function in the Year 2 cohort were uniformly comparable to those in the ITT population at 6 and 12 months [–0.56 (n = 113) vs –0.37 (n = 111)], and were maintained over 24 months of treatment. A majority of patients (86% LEF, 82% SSZ) reported changes in physical function that met or exceeded MCID.

**MN302/304 results.** In the Year 2 cohort, mean changes in HAQ-DI from baseline to Months 6, 12, and 24 were ~0.43, ~0.56, and ~0.48 in LEF vs ~0.40, ~0.61, and ~0.56 in MTX treatment groups, respectively; they were statistically equivalent at Month 24 (Table 2, Figure 2). Improvements in physical function at 12 months in the Year 2 cohort were uniformly comparable to those in the ITT population at 6 and 12 months [–0.56 (n = 134) vs –0.37 (n = 134)], and were maintained over 24 months of treatment. Mean changes from baseline considerably exceeded MCID, indicating a large majority of patients in the Year 2 cohorts achieved clinically important improvement.

Mean and median changes from baseline to 24 months in HAQ Disability Index in Year 2 cohorts.

<table>
<thead>
<tr>
<th></th>
<th>LEF</th>
<th>MTX</th>
<th>SSZ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>US301</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>97</td>
<td>101</td>
<td>—</td>
</tr>
<tr>
<td>Baseline mean</td>
<td>1.2</td>
<td>1.2</td>
<td>—</td>
</tr>
<tr>
<td>Mean (median) change at 6 mo</td>
<td>–0.55 (–0.50)</td>
<td>–0.34 (–0.25)</td>
<td>—</td>
</tr>
<tr>
<td>Mean (median) change at 12 mo</td>
<td>–0.60 (–0.63)</td>
<td>–0.37 (–0.38)</td>
<td>—</td>
</tr>
<tr>
<td>Mean (median) change at 24 mo</td>
<td>–0.60* (–0.63)</td>
<td>–0.37 (–0.38)</td>
<td>—</td>
</tr>
<tr>
<td>Mean change at 24 mo, %</td>
<td>50</td>
<td>31</td>
<td>—</td>
</tr>
<tr>
<td><strong>MN301/303/305</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>51</td>
<td>—</td>
<td>46</td>
</tr>
<tr>
<td>Baseline mean</td>
<td>1.6</td>
<td>—</td>
<td>1.5</td>
</tr>
<tr>
<td>Mean (median) change at 6 mo</td>
<td>–0.75 (–0.75)</td>
<td>—</td>
<td>–0.61 (–0.50)</td>
</tr>
<tr>
<td>Mean (median) change at 12 mo</td>
<td>–0.70 (–0.63)</td>
<td>—</td>
<td>–0.62 (–0.57)</td>
</tr>
<tr>
<td>Mean (median) change at 24 mo</td>
<td>–0.73 (–0.69)</td>
<td>—</td>
<td>–0.56 (–0.38)</td>
</tr>
<tr>
<td>Mean change at 24 mo, %</td>
<td>46</td>
<td>—</td>
<td>37</td>
</tr>
<tr>
<td><strong>MN302/304</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of patients</td>
<td>248</td>
<td>273</td>
<td>—</td>
</tr>
<tr>
<td>Baseline mean</td>
<td>1.5</td>
<td>1.5</td>
<td>—</td>
</tr>
<tr>
<td>Mean (median) change at 6 mo</td>
<td>–0.43 (–0.38)</td>
<td>–0.40 (–0.38)</td>
<td>—</td>
</tr>
<tr>
<td>Mean (median) change at 12 mo</td>
<td>–0.56 (–0.50)</td>
<td>–0.61 (–0.56)</td>
<td>—</td>
</tr>
<tr>
<td>Mean (median) change at 24 mo</td>
<td>–0.48 (–0.38)</td>
<td>–0.56 (–0.50)</td>
<td>—</td>
</tr>
<tr>
<td>Mean change at 24 mo, %</td>
<td>32</td>
<td>37</td>
<td>—</td>
</tr>
</tbody>
</table>

*p = 0.005, LEF vs MTX.

**DISCUSSION**

Analyses of 3 phase III RCT presented here compared 24 months of blinded treatment in patients receiving LEF, MTX, and SSZ, providing strong, consistent data that improvements in physical function and HRQOL are sustained from 12 to 24 months in patients who continue therapy. Mean changes from baseline considerably exceeded MCID, indicating a large majority of patients in the Year 2 cohorts achieved clinically important improvement.

In MN301/303/305, after 24 months’ blinded treatment, numerical differences in changes in HAQ-DI scores between LEF and SSZ patients approached 0.22 units (Figure 4); lack of statistical significance may be attributable to small sample sizes in the Year 2 cohort. Compared with mean and median scores at Month 12, statistical decreases in HAQ-DI at Month 24 in the LEF treatment group in MN302/304 were small (0.08 to 0.12), and likely do not reflect clinically meaningful differences, as evidenced by the 95% confidence interval around mean change in HAQ-DI (0.029, 0.140), which does not include 0.22 (Table 2). Measures of physical function and HRQOL reflect disease progression over time, even in patients considered to be “doing well” while receiving standard of care, including DMARDs. Data published prior to introduction of the new
Figure 2. HAQ-DI results over time. Improvements in HAQ-DI were maximal at 6 months and sustained over 24 months in all 3 studies. Improvement with leflunomide (LEF) was statistically better than methotrexate (MTX) in US301 (p = 0.005) and equivalent in MN302/4. SSZ: sulphasalazine.
DMARD reported stabilization, but not improvement, in HAQ scores with standard of care\textsuperscript{45,46}. The phase III RCT reported here were the first to show that DMARD therapy can result in significant improvements in physical function, with maximal benefit at 6 months, sustained over 12–24 months’ treatment, as confirmed in the subsequent Anti-TNF Trial in RA with Concomitant Therapy (ATTRACT), Etanercept in Early RA (ERA), and Trial of Etanercept and Methotrexate with Radiographic Patient Outcomes (TEMPO) trials\textsuperscript{15,18,21,23-25,47-49}. Importantly, the blind appears to have been successfully maintained in these phase III trials, for as long as 36 months in the US301 extension\textsuperscript{5}.

Protocol US301 first demonstrated that improvements in SF-36 reflected treatment effects in active RA, in mental as well as physical domains. Subsequent data from RCT with biologic agents have confirmed correlations between HAQ and SF-36\textsuperscript{14,17,50} (below).

Despite significant decrements at baseline, 12–24 month treatment with LEF or MTX resulted in mean values in SF-36 domains that approached age and sex matched US norms, with the exception of the domains with lowest baseline values: physical function and role-physical (Figure 5). Although SF-36 was not included in the MN trials, changes in HAQ-DI observed in both were similar to those reported

---

**Figure 3.** Changes in the SF-36 domains at endpoint in the US301 Year 2 cohort. *Improvements with leflunomide (LEF) were statistically greater (p < 0.05) than with methotrexate (MTX) in these domains. Improvements in all domain scores with active treatment met or exceeded MCID, indicated by broken lines at 5–10 points.

**Figure 4.** Changes in the HAQ-DI at endpoint in the Year 2 cohorts. Improvements in the HAQ-DI considerably exceeded MCID in all 3 studies, represented by the broken line at −0.22.
in US301, reflecting improvements in the mental as well as the physical domains of HRQOL of a magnitude considered to be clinically meaningful.

In the 3 phase III RCT, 30% to 40% of ITT and Year 2 cohort patients reported disease durations ≤ 2 years and/or were DMARD-naive. These results indicate clinically meaningful responses in patients with early disease receiving monotherapy, similar to data reported with combination and biologic therapies. Even in patients with early disease, HAQ scores predict loss of work and premature mortality, correlating with progression to disability and increased cost of illness. 

Findings from these trials have been criticized because of dropout rates between the ITT and Year 2 cohort populations. US301 was the first 24-month RCT conducted in RA offering rescue on or after 4 months for nonresponders in the placebo and active treatment groups — “incentive” in a sense for patients to switch to alternative therapy. In this pure placebo controlled trial, overall completion rates were 65%–72%; 52%–53% of the ITT population entered the second year receiving their originally assigned treatment. In comparison, in MN302, an active controlled trial, one-year completion rates were 70%–78%; 58%–64% chose to continue treatment in MN304. In the MN302 ITT population, 51%–56% completed 24 months’ treatment, which compares favorably with completion rates of 52%–59% in MTX treatment groups in the ERA and TEMPO trials, respectively, and 55%–68% across all infliximab arms of the ATTRACT, with background MTX therapy. As with the ATTRACT trial, sensitivity analyses of HAQ and SF-36 data were performed to account for missing data, and confirmed the robustness of the findings.

More importantly, treatment-associated improvements in HAQ and SF-36 reported here compare favorably with results reported from the ATTRACT in patients with 7–10 years’ disease duration, as well as the ERA and the Active controlled Study of Patients receiving Infliximab for RA of Early onset (ASPIRE) in early RA. At baseline, 100% of patients enrolled in ATTRACT reported role-phys-

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study</th>
<th>Scales</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kavanaugh14</td>
<td>Infliximab/ATTRACT</td>
<td>SF-36 PCS PF</td>
<td>−0.51</td>
</tr>
<tr>
<td>Kosinski17</td>
<td>Etanercept/ERA</td>
<td>SF-36 PCS PF</td>
<td>−0.60</td>
</tr>
<tr>
<td>Lubeck50</td>
<td>Etanercept/RAPOLO</td>
<td>SF-36 PCF PF</td>
<td>−0.79</td>
</tr>
<tr>
<td>Strand</td>
<td>Leflunomide/US301</td>
<td>SF-36 PCS PF</td>
<td>−0.60</td>
</tr>
</tbody>
</table>

PF: Physical function domain of SF-36.
ical scores of 0, compared with 49% in US301 and none in ERA; mean baseline scores were 0.0, 22.4, and 30.9, respectively, indicating a significant influence of active RA on HRQOL. Improvements ranged from medians of 12.7–26.9 with infliximab + MTX and means of 37.6 with LEF and 26.1 with MTX over 24 months’ treatment compared with means of 9.9 with MTX and 10.8 with etanercept over 12 months. Baseline physical component scores were similar across trials: 23.9–25.8 in ATTRACT, 30.2–30.9 in US301, and 28.0–29.2 in ERA; improvements with active treatment ranged from 4.6 to 6.9 (median 24 mo) with infliximab + MTX to 9.6 and 10.7 (mean 12 mo), respectively, with MTX and etanercept, comparable with values reported in US301 at 12 and 24 months (Table 3). In each trial population, baseline scores were more than 2 standard deviations below the US norm; posttreatment values increased to within 1–2 SD of normative scores of 50. In view of changes reported in all 8 domains of the SF-36, it is clear that physical component summary scores alone do not reflect the full range of improvement in HRQOL that occurs when treatments positively affect physical function in patients with active RA. As with the HAQ-DI, treatment-associated improvements in HRQOL were maximal within 6 months and were sustained over 12 to 24 months.

Singh, et al determined that a 1-point increase in HAQ-DI over the first 2 years of disease results in 90% greater disability and 87% higher costs over the next 3 years, and 75% greater disability and 74% higher costs over the next 8 years. In addition, Yelin and Wanke found that RA patients in the most affected quartile of physical function incur total annual and hospital costs 2.55 and 6.97 times higher, respectively, than those incurred by patients in the least affected quartile.

Figure 5. Final SF-36 domain scores in the US301 Year 2 cohort. Two lines indicate baseline values prior to treatment, the third portrays US norms for an age and sex matched population. Treatment-associated improvements at 12 and 24 months approach US norms for all domains except physical functioning and role physical.
REFERENCES

7. Strand V, Pincus T. The Health Assessment Questionnaire provides a single effective measure to discriminate active from placebo treatment in randomized trials in rheumatoid arthritis [abstract]. Arthritis Rheum 2003;48 Suppl 3656.
19. Strand V, Crawford B, Maetzler A, Bombardier B, Tugwell P. A cost and consequences comparison of therapy with methotrexate or leflunomide in patients with active rheumatoid arthritis. [submitted for publication].
31. Ware JE Jr, Kosinski M, Keller S. SF-36 physical and mental health summary scales: A user’s manual. Boston: The Health Institute,


