ARTHRITIS & RHEUMATISM
Vol. 42, No. 9, September 1999, pp 1870–1878

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

VIBEKE STRAND, PETER TUGWELL, CLAIRE BOMBARDIER, ANDREAS MAETZEL, BRUCE CRAWFORD, CATHERINE DORRIER, ANN THOMPSON, and GEORGE WELLS, on behalf of the LEFLUNOMIDE RHEUMATOID ARTHRITIS INVESTIGATORS GROUP

Objective. **To assess the efficacy of leflunomide or methotrexate compared with placebo in improving function and health-related quality of life in patients with active rheumatoid arthritis (RA), and to examine correlations between response status (as defined by the American College of Rheumatology [ACR] response criteria) and improvement in these measures.**

Methods. **This 52-week, multicenter, double-blind, controlled trial compared responses to the Health Assessment Questionnaire (HAQ), modified Health Assessment Questionnaire (MHAQ), Problem Elicitation Technique (PET), Medical Outcomes Study Short Form 36 (SF-36), and questions regarding work productivity among 3 treatment groups (leflunomide, methotrexate, and placebo). Improvement in the PET top 5 and SF-36 scales and component scores were compared with ACR response rates.**

Results. **Clinically meaningful and statistically significant ($P < 0.0001$) improvement in measures of function and heath-related quality of life (MHAQ scores, all scales and disability index of the HAQ, weighted top 5 score of the PET, 5 of 8 scales and physical component score of the SF-36, and work productivity) was seen during treatment with leflunomide in comparison with placebo. Methotrexate administration resulted in significant improvements ($P < 0.05$) in comparison with placebo in the MHAQ scores, HAQ disability index, weighted top 5 score of the PET, physical component score of the SF-36, and bodily pain scale. Compared with methotrexate, leflunomide administration resulted in significantly ($P < 0.01$) more improvement in the MHAQ scores, 5 of 8 scales and disability index of the HAQ, weighted top 5 score of the PET, and 2 of 8 scales and physical component score of the SF-36. Improvements in the PET score, SF-36 physical component score, and work productivity correlated with the ACR responder rates of ≥20% and ≥50% improvement.**

Conclusion. **Significant improvements in function and health-related quality of life occurred in patients with RA during treatment with leflunomide or methotrexate. These findings were clinically meaningful and correlated with the ACR response status.**

Rheumatoid arthritis (RA) is characterized by a symmetric, erosive synovitis, which is usually progressive despite treatment. The resulting joint deformity and loss of function lead to disability and deterioration in health-related quality of life (1). Patients with active RA complain that they are most affected by pain and loss of function when performing regular activities.

Several disease-specific instruments have been developed to assess functional status in RA. The Health Assessment Questionnaire (HAQ), modified Health As-

Supported by Hoechst Marion Roussel (HMR).

Vibeke Strand, MD: Stanford University, San Francisco, California; Peter Tugwell, MD, George Wells, PhD: University of Ottawa, Ottawa, Canada; Claire Bombardier, MD, Andreas Maetzel, MD, PhD: University of Toronto, Ontario, Canada; Bruce Crawford, MA: MAPI Values, Boston, Massachusetts; Catherine Dorrier, MS: Quintiles Transnational, Arlington, VA; Ann Thompson, MBA, BSN: Hoechst Marion Roussel, Kansas City, Missouri.

Dr. Strand serves as a consultant to HMR in clinical and regulatory affairs. Drs. Tugwell, Bombardier, Maetzel, and Wells are consultants to HMR, Canada, related to the pharmacoeconomic analysis of this protocol. Dr. Tugwell was also a member of the Data Safety Monitoring Board for this protocol. Mr. Crawford serves as a consultant to HMR. Ms Dorrier is a statistical consultant to HMR.

Address reprint requests to Ann K. Thompson, BSN, MBA, Hoechst Marion Roussel, 10326 Marion Park Drive, Kansas City, MO 64134-0627.

Submitted for publication December 28, 1998; accepted in revised form April 27, 1999.

sessment Questionnaire (MHAQ), and Problem Elicitation Technique (PET) have been validated in randomized controlled clinical trials, and responses to these instruments can reflect impairment in performance of daily and other essential activities (2–4). The PET asks patients to rank the activities included in the HAQ as to their importance and the degree to which they are affected by the disease. Its use in a randomized controlled trial of cyclosporine was reported in 1992 (5). All 3 measures have been shown to be responsive to change in disease status and to correlate with the American College of Rheumatology (ACR) response criteria (4,6).

To date, few randomized controlled trials in RA have assessed the effect of treatment on improving or stabilizing physical function and/or health-related quality of life. Bombardier et al showed that treatment with auranofin over 6 months was associated with significantly greater improvement in function, as measured by the HAQ, when compared with placebo (7). Tugwell et al reported statistically significant improvements in health status with methotrexate treatment compared with placebo over 6 months, as measured by the Lee Functional Index, the McMaster Toronto Arthritis Patient Disability instrument, and the McMaster Health Index Questionnaire (8). A subsequent report by Bombardier and colleagues showed that a 6-month administration of cyclosporine was superior to placebo in improving function according to the PET, HAQ, and Arthritis Impact Measurement Scales (5).

The Medical Outcomes Study Short Form 36 (SF-36) is a generic instrument that has been validated in normal and diseased populations for assessing health-related quality of life (9). It was first used in a clinical trial of minocycline for patients with RA in 1997 (10). Although the changes observed correlated with the results as assessed by the MHAQ and patient and physician global assessments, there were no significant differences in any measures between the active treatment and placebo groups after 48 weeks of treatment. In an observational study of 137 RA patients over 4 months, Talamo et al confirmed the close correlation between the physical functioning scale of the SF-36 and the HAQ (11). Ruta and colleagues assessed the responsiveness to change of the SF-36 in 240 British RA patients observed over 3 months, comparing it with the ACR response criteria, including the MHAQ (12). Baseline scale and physical and mental component scores correlated with the ACR functional class. The SF-36 was similarly assessed by Wells et al in a multicenter controlled trial comparing generic quality of life instruments in 40 patients initiating methotrexate therapy, who were examined at baseline and 3 months (13). Although the SF-36 physical component score was not as sensitive to change as the Nottingham Health Profile and Rheumatoid Arthritis Quality of Life following 6 months of treatment, it showed similar positive (57.1%) and negative (83.3%) agreements with the ACR ≥20% response criteria.

Leflunomide (Arava; Hoechst Marion Roussel, Kansas City, MO) is a novel immunomodulatory agent with prophylactic and therapeutic effects in animal models of autoimmune disease (14). It was first demonstrated to be effective in a 6-month placebo-controlled phase II study in 402 patients with active RA (15), and was subsequently studied in 3 randomized controlled phase III trials of 6 and 12 months' duration conducted in the US, Europe, South Africa, Australia, and New Zealand. Clinical data have shown it to be an effective disease-modifying antirheumatic drug (DMARD) therapy, equivalent to methotrexate and sulfasalazine, which retards disease progression as measured on radiographs (16–18). This article describes the effects of leflunomide administration, as compared with placebo and methotrexate, in improving function and health-related quality of life in a 12-month placebo-controlled trial conducted in the US and Canada in 482 methotrexate-naïve patients with active disease. Clinical results from this trial have been reported elsewhere (19).

PATIENTS AND METHODS

This study was a 12-month, randomized, multicenter, double-blind, placebo-controlled trial to assess the safety and efficacy of leflunomide treatment compared with placebo and methotrexate. The study was approved by appropriate Institutional Review Boards before patients were enrolled. Men and women 18 years of age or older were eligible for treatment if they met the ACR (formerly, the American Rheumatism Association) criteria for RA (20) and had RA for ≥6 months. Active disease was defined by the presence of 3 of the following 4 criteria: ≥9 tender joints, ≥6 swollen joints, morning stiffness of ≥45 minutes, and a Westergren erythrocyte sedimentation rate (ESR) of ≥28 mm/hour. Notably, to be eligible, patients could not have previously received methotrexate, and treatment with all other DMARDs must have been discontinued for at least 30 days. Prednisone ≤10 mg/day (or the equivalent) and nonsteroidal antiinflammatory drugs were permitted if doses had been stable for at least 30 days before enrollment and remained so during treatment.

Patients were randomly assigned in a 3:2:3 distribution to 1 of 3 treatment groups: leflunomide at 20 mg daily, placebo, or methotrexate at 7.5–15.0 mg weekly. To preserve the study blind, all patients received once daily oral leflunomide or matching leflunomide placebo, and once weekly oral methotrexate or matching methotrexate placebo. A 100-mg loading dose of leflunomide or leflunomide placebo was administered

for the first 3 days to allow a steady-state plasma concentration to be reached within 6–8 weeks. If active disease (as defined above) was still present at the sixth week of treatment, the dose of methotrexate or methotrexate placebo was mandated to be increased to 15.0 mg over weeks 7 to 9 and continued thereafter.

A total of 482 patients were enrolled, and 480 (182 leflunomide, 118 placebo, and 180 methotrexate) were evaluable for clinical response using a modified intention-to-treat analysis, which was defined as inclusion of all patients who received at least 1 dose of study drug with at least 1 followup visit. This report presents a secondary analysis of 438 patients (166 leflunomide, 102 placebo, and 170 methotrexate), all of whom completed baseline and 1 or more followup HAQ and SF-36 questionnaires. This population included 438 patients instead of 480 because 4 subjects did not complete a baseline questionnaire (a validated Spanish translation of the SF-36 was lacking at the time of this study's initiation) and 20 subjects exited early without completing followup questionnaires. Eighteen questionnaires were excluded due to inconsistent responses, 9 at baseline and 9 at followup, as calculated by the "response consistency index" developed by The Health Institute (21). The demographic and baseline clinical characteristics of those patients who did not complete the questionnaires were similar to those of the entire protocol population.

The following measures were collected at baseline and at each monthly visit as components of the ACR response, which was the primary outcome measure: tender and swollen joint counts (28 joints), patient and physician global assessments of disease activity (on a 0–100-mm visual analog scale [VAS]), ratings on a pain intensity scale (0–100-mm VAS), MHAQ scores, Westergren ESR, and C-reactive protein levels. The MHAQ assessment is a single page of 8 questions (a subset of the 20 questions included in the full HAQ described below) about functional activities performed on a daily basis, and each item is scored by patients on a scale from 0 (without difficulty) to 3 (unable to do).

Physical function and health-related quality of life were assessed across the 3 treatments by calculating the mean changes from baseline to end point or to study withdrawal, in each of the following measures: HAQ, PET, SF-36, and questions related to work productivity. The HAQ, a disease-specific instrument, includes 20 questions divided into 8 functional categories of 2 or 3 questions each, concerning activities performed on a daily basis (dressing and grooming, arising, walking, eating, hygiene, reaching, gripping, and activities). Each question is scored 0 (without difficulty) to 3 (unable to do). The highest scores in each of the categories are then summed (range 0–24) and divided by the number of categories scored, to give a disability index that ranges from 0 to 3, with higher scores indicating more disability. Mean changes in each functional category are also reported, as well as an unweighted sum of the means in each category, divided by the number of categories (22).

The PET is also a disease-specific instrument which asks patients first to identify functional activities (as enumerated by the HAQ in this protocol) that are most affected by their RA and that they would most like to see improved by treatment. Patients then rank the difficulty, severity, and/or frequency of performing these activities on a 7-point scale (e.g., 0 = no difficulty to 7 = unable to do). They are then asked to rate the level of importance of each problem on a 7-point scale (0 = not at all important to 7 = most important). The weighted top 5 score of the PET is determined by summing the scores for the 5 most important problems.

Table 1. Comparison of physical functions assessed by the HAQ, MHAQ, and SF-36*

Activities assessed	HAQ	MHAQ	SF-36 physical functioning
Walking	+	+	+
Climbing steps	+		+
Reaching	+	+	
Getting in and out of car	+	+	
Arising	+		
Reaching over head	+		
Gripping	+	+	
Eating	+	+	+
Self care ADL			
Hygiene	+	+	
Dressing, grooming	+	+	+
Instrumental activities	+		
Discretionary activities			
Walking >1 mile			+
Climbing several sets of stairs			+
Moderate activities			+
Vigorous activities			+

* HAQ = Health Assessment Questionnaire; MHAQ = modified Health Assessment Questionnaire; SF-36 = Short Form 36; ADL = activities of daily living.

The SF-36 is a generic instrument with scores that are based on responses to individual questions, which are summarized into 8 scales, each of which measures a health concept. These scales include function domains and aspects of well-being, as follows: physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. The physical functioning scale includes both essential and discretionary activities. For each of the SF-36 scales, necessary items are recoded so that higher values indicate better health, and then summed. The summed scores are transformed to a 0–100 scale following its designated scoring algorithm, with higher scores reflecting better quality of life. These 8 scales, weighted according to normative data, are also combined into summary physical and mental component scores, which, again, are scored from 0 to 100, with higher scores reflecting better quality of life. Table 1 presents the questions about daily and discretionary functions that are common to each of these instruments.

Questions regarding work productivity were abstracted from the Work Limitations Questionnaire of the National Opinion Research Center survey. Patients were asked to rank their level of difficulty with work-related activities due to health problems and health concerns on a 6-point scale (e.g., 1 = "none" to 6 = "not done, can't do it"). An additional response category was added for "not part of job." The score for work productivity follows the same algorithm as the SF-36 scales. If the patient responded "not part of job," that question was removed from their score. Scores were presented on a 0–100 scale, with higher scores reflecting higher productivity at home, school, and work.

Table 2. Demographics and disease characteristics of the study patients at baseline*

	Leflunomide (n = 182)	Placebo (n = 118)	Methotrexate (n = 182)
Sex, % female	73	70	75
Age, years (mean ± SD)	54.1 ± 12.0	54.6 ± 10.7	53.3 ± 11.8
Disease duration, years (mean ± SD)	7.0 ± 8.6	6.9 ± 8.0	6.5 ± 8.1
% with disease ≤2 years	39	33	40
% with disease >5 years	44	45	37
% rheumatoid factor positive	65	60	59
Mean ± SD number DMARDs failed	0.8 ± 1.0	0.9 ± 0.9	0.9 ± 1.0
% with no prior DMARD treatment	46	40	44
% taking concomitant NSAIDs	75	65	70
% taking concomitant steroids	54	55	53
% ACR functional class I	13	9	12
% ACR functional class II	74	75	77
% ACR functional class III	12	14	11
Total % ACR functional class I–III	99	100	100
% ACR functional class IV	1.1	0	0
% with erosions on baseline radiograph	70	78	66

* $P > 0.05$ for all baseline comparisons. DMARDs = disease-modifying antirheumatic drugs; NSAIDs = nonsteroidal antiinflammatory drugs; ACR = American College of Rheumatology.

Statistical analysis. All analyses were performed on the intention-to-treat patient population. Comparisons of baseline and posttreatment values across the 3 treatment groups used chi-square analysis for categorical variables and analysis of variance for continuous variables. The end point of analysis presented herein was 12 months. When necessary, analyses used the last observation carried forward, with all function and health-related quality of life measures required only at month 6 and month 12, or at early termination from the study. Because data from the health-related quality of life instruments were not normally distributed at baseline, the Van Elteren extension to the Wilcoxon rank sum test was applied to continuous variables and the Cochran-Mantel-Haenszel chi-square test for categorical variables, for comparing baseline characteristics.

The primary analysis for treatment effect consisted of 2 steps. First, a multivariate analysis of covariance omnibus test was used. If significant differences were found, individual scale scores were examined. The multivariate omnibus test was initially performed to try to minimize the potential for an inflated alpha risk. Since the omnibus tests were statistically significant for differences at study exit, thus signifying an overall difference between groups, univariate tests on individual scales were performed. If there was no statistical difference between treatments using the omnibus test, individual scale scores were not examined. Individual scale scores were analyzed using a general linear model including treatment, region, treatment by region, and any differences identified at baseline as factors.

Comparisons of ≥20% and ≥50% improvement in the PET weighted top 5 scores (a disease-specific measure), SF-36 physical functioning and bodily pain scales, and physical and mental component scores (a generic measure), in responders and nonresponders (as defined by the ACR criteria of ≥20% and ≥50% improvement) were examined. Because the modified HAQ is a component of the ACR criteria and it, in turn, is derived from the HAQ, comparisons of the results from these measures with the ACR response would have been confounded and so were not made. To understand correlations between the ACR composite index and changes in function and health-related quality of life measures, scatterplots of individual patients' changes in tender joint count versus PET weighted top 5 score and SF-36 physical component score for responders and nonresponders by ACR ≥20% response criteria were generated.

RESULTS

There were no significant differences between treatment groups in the demographic or baseline disease characteristics at study entry (Table 2). The overall mean HAQ disability index across the 3 treatment groups was 1.3, and the overall mean MHAQ score was 0.78–0.89. Baseline SF-36 scales reflected values that were significantly lower than US norms (Figure 1). The baseline HAQ disability index, PET weighted top 5, and SF-36 scores were consistent with the ACR functional class at study entry (Table 3).

The weekly dose of methotrexate was increased to 15 mg in 109 (60%) of the patients who were receiving methotrexate. Increased doses of methotrexate placebo were mandated in 95 (52%) of the leflunomide patients and 81 (69%) of the placebo patients. Few dose reductions were necessary: 3 (2%) in the leflunomide group, none in the placebo group, and 4 (2%) in methotrexate-treated patients.

Because the increase in the weekly dose of methotrexate from 7.5 to 15 mg could occur only in patients who failed to respond, ACR response rates in those patients were not different from the response rates in patients whose dose remained at 7.5 mg: 45% and 47%,

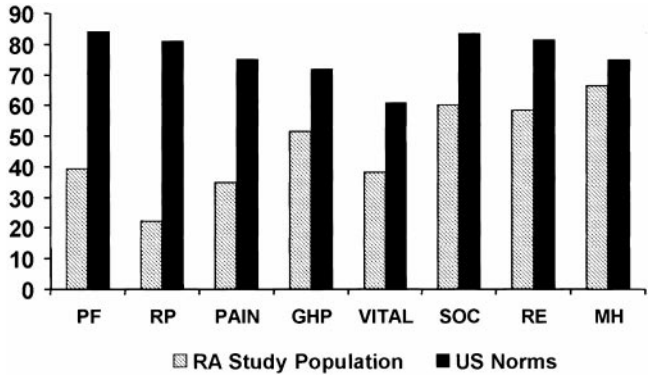

Figure 1. Mean results on the Short Form 36 scales at baseline in rheumatoid arthritis (RA) patients compared with US norms. Because baseline characteristics were equally distributed between treatment groups, the entire protocol population was compared with US norms. PF = physical functioning; RP = role-physical; PAIN = bodily pain; GHP = general health profile; VITAL = vitality; SOC = social functioning; RE = role-emotional; MH = mental health.

respectively. This was also true in the leflunomide and placebo treatment groups, in which ACR response rates were 51% in the leflunomide group and 25% in the placebo group among patients whose dose of methotrexate placebo was increased, compared with 53% and 27%, respectively, in patients whose placebo dose remained the same.

As shown in Table 4, significantly greater improvements in function and health-related quality of life were reported with leflunomide treatment when compared with placebo as measured by the MHAQ, all scales and disability index of the HAQ, weighted top 5 score of the PET, 5 of 8 scales (physical functioning, bodily pain, general health profile, vitality, and social functioning) and the physical component score of the SF-36, and work productivity (all $P < 0.0001$). The relative percentage improvement compared with placebo in each of the parameters is also presented in Table

Table 3. HAQ/PET/SF-36 scores compared with ACR functional class at baseline

	Class I	Class II	Class III	Class IV
SF-36 physical functioning (0–100)				
No. (total = 427)	53	324	48	2
Mean score	51.9	40.0	22.3	0.0
PET weighted difficulty/top 5 (0–100)				
No. (total = 436)	54	330	50	2
Mean score	18.3	20.4	28.6	41.0
HAQ disability index (0–3)				
No. (total = 435)	54	329	50	2
Mean score	1.1	1.3	1.8	2.4

* PET = Problem Elicitation Technique (see Tables 1 and 2 for other definitions).

Table 4. Mean changes in function and health-related quality of life measures at end point and 6 months versus baseline (intention-to-treat analysis)*

	Leflunomide	Placebo	Methotrexate
HAQ disability index			
n	166	101	169
Baseline	1.30	1.31	1.30
Mean change at 6 months	−0.46	0.07	−0.30
Mean change at end point	−0.45†‡	0.03	−0.26§
% change from baseline	35	−2	20
Mean % change vs. placebo	37	–	22
MHAQ			
n	182	118	180
Baseline	0.78	0.89	0.79
Mean change at 6 months	−0.32	0.03	−0.17
Mean change at end point	−0.29†‡	0.07	−0.15§
% change from baseline	37	−8	19
Mean % change vs. placebo	45	–	27
PET top 5			
n	166	101	170
Baseline	21.2	22.4	20.4
Mean change at 6 months	−6.7	−1.1	−3.1
Mean change at end point	−6.9†‡	−0.66	−3.4§
% change from baseline	35	3	17
Mean % change vs. placebo	29	–	13
SF-36 physical component			
n	157	101	162
Baseline	30.0	28.9	29.7
Mean change at 6 months	7.3	0.9	4.9
Mean change at end point	7.6†‡	1.0	4.6§
% change from baseline	25	3	15
Mean % change vs. placebo	22	–	12
SF-36 mental component			
n	157	101	162
Baseline	46.8	48.3	48.5
Mean change at 6 months	3.0	1.0	1.5
Mean change at end point	1.5	0.8	0.9
% change from baseline	3	2	2
Mean % change vs. placebo	1	–	0
Work productivity			
n	138	92	148
Baseline	53.3	52.9	51.9
Mean change at 6 months	11.1	3.3	7.7
Mean change at end point	9.8†	0.3	7.5§
% change from baseline	18	0.5	14
Mean % change vs. placebo	18	–	14

* End point is 12-month value or, if missing, last observation carried forward. See Tables 1 and 3 for definitions.
† $P < 0.0001$ versus placebo.
‡ $P < 0.01$ versus methotrexate.
§ $P < 0.05$ versus placebo.

Table 5. Improvement of ≥20% and ≥50% in function and health-related quality of life measures at end point versus baseline (intention-to-treat analysis)*

	Leflunomide	Placebo	Methotrexate
ACR responder			
≥20% responder	52	26	46
≥50% responder	34	8	23
HAQ disability index			
≥20% improvement	57	27	49
≥50% improvement	42	14	30
MHAQ			
≥20% improvement	64	35	53
≥50% improvement	53	18	37
PET top 5			
≥20% improvement	58	31	48
≥50% improvement	35	11	23
SF-36 physical component			
≥20% improvement	49	27	39
≥50% improvement	30	5	23
SF-36 mental component			
≥20% improvement	26	21	26
≥50% improvement	7	4	7
Work productivity			
≥20% improvement	39	20	38
≥50% improvement	22	10	23

* End point is 12-month value or, if missing, last observation carried forward. Values are the percentage of patients. See Tables 1 and 3 for definitions.

4. The HAQ disability index decreased by –0.45, from 1.30 to 0.85, a 37% improvement relative to placebo. The MHAQ score improved by –0.29, from 0.78 to 0.49, a 45% improvement relative to placebo. The PET top 5 score improved by –6.9, from 21.2 at baseline, a 29% improvement relative to placebo. The physical component score of the SF-36 increased by 7.6 points, from 30.0, a mean change of 22% compared with placebo, whereas the mental component score showed only a small change. As shown in Table 5, the percentage of patients with improvements of ≥20% and ≥50% in measures of function and the physical component score of the SF-36 were similar to (or exceeded) the ≥20% and ≥50% ACR response rates.

In comparison with placebo, methotrexate administration resulted in significant improvements in the MHAQ scores, HAQ disability index, weighted top 5 score of the PET, physical component score of the SF-36, and bodily pain (all $P < 0.05$) (Table 4 and Figures 2 and 3). The HAQ disability index decreased by –0.26, from 1.30 to 1.04, a 22% improvement relative to placebo. The MHAQ score improved by –0.15, from 0.79 to 0.64, a 27% improvement relative to placebo. The PET top 5 score improved by –3.4, from 20.4 at

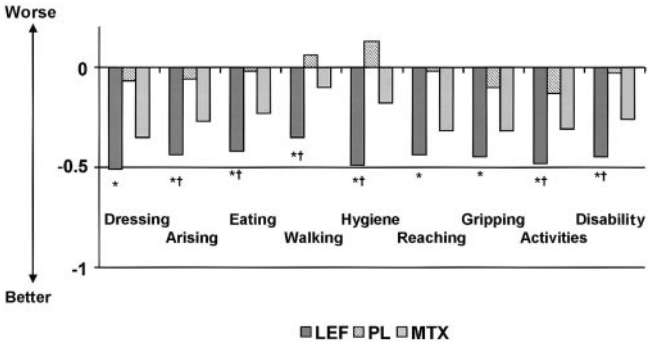

Figure 2. Improvement in Health Assessment Questionnaire categories. Mean improvement in each category, calculated as the end point minus the baseline score, is presented for each treatment group (intention-to-treat population). LEF = leflunomide; PL = placebo; MTX = methotrexate. * = $P < 0.05$ LEF versus PL; † = $P < 0.05$ LEF versus MTX.

baseline, a 13% improvement relative to placebo. The physical component score of the SF-36 increased by 4.6 points, from 29.7, a mean change of 12% compared with placebo, whereas the mental component score showed no change. As with leflunomide treatment, improvements in measures of function and the physical component score of the SF-36 were similar to (or exceeded) the ≥20% and ≥50% ACR response rates (Table 5).

Figures 2 and 3 display changes in the 8 categories of the HAQ and 8 scales of the SF-36 in each treatment group. These changes, evident at 6 months and sustained over 12 months of treatment (data not shown), are presented as the mean change from baseline using an intention-to-treat analysis. Deterioration or no change in function (as measured by the MHAQ) was evident in the placebo population over this 12-month period.

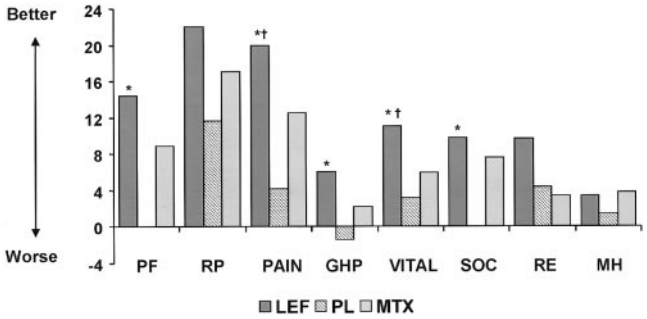

Figure 3. Improvement in Short Form 36 scales. Mean improvement in each scale, calculated as the end point minus the baseline scores, is presented for each treatment group (intention-to-treat population). * = $P < 0.05$ LEF versus PL; † = $P < 0.05$ LEF versus MTX. See Figures 1 and 2 for definitions.

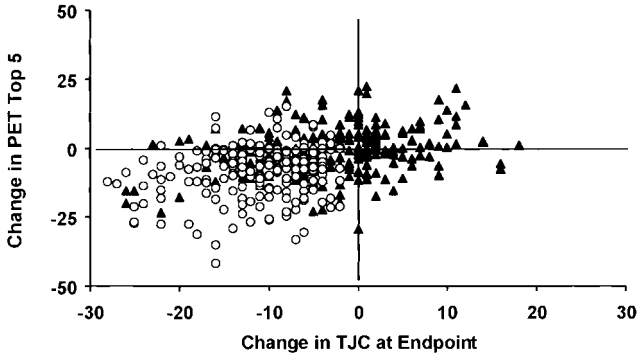

Figure 4. Change in Problem Elicitation Technique (PET) top 5 scores at end point versus American College of Rheumatology (ACR) responders by total joint count (TJC) (leflunomide treatment group, intention-to-treat population). ACR responders (≥20% improvement) are plotted as open circles, and nonresponders as dark triangles. Improvement corresponds to the lower left hand field, indicating corresponding decreases in tender joint count and PET top 5 scores.

In comparison with methotrexate, leflunomide administration resulted in significantly greater improvements in the MHAQ scores ($P \leq 0.01$), 5 of 8 scales and the disability index of the HAQ ($P \leq 0.01$), weighted top 5 score of the PET ($P \leq 0.001$), and 2 of 8 scales (bodily pain and vitality) and the physical component score of the SF-36 ($P \leq 0.01$) (Table 4 and Figures 2 and 3).

Significant concordance between patients who were ACR responders (≥20% and ≥50%) and those who demonstrated ≥20% and ≥50% improvement in the disease-specific and generic measures of function and health-related quality of life was evident. Figures 4 and 5 present scatterplots of change in tender joint count versus change in PET weighted top 5 and SF-36 physical component scores in patients receiving leflunomide who were ACR ≥20% responders.

DISCUSSION

In previous clinical trials of shorter duration (18 weeks to 9 months), retrospectively applied ACR response rates of 40.3% (compared with 8.4% in those receiving placebo) and 64.7% (compared with 28.8% in those receiving auranofin) have been reported following methotrexate administration. Recently, an ACR response rate of 39% for methotrexate, compared with 12% for placebo, was reported at end point in a 6-month trial, which also examined cyclosporine treatment. Response rates following methotrexate treatment in the present trial compare favorably with these published results.

In this double-blind, placebo-controlled, multicenter trial comparing leflunomide, methotrexate, and placebo in 482 patients with active RA, detailed analyses of disease-specific and generic measures demonstrated that leflunomide and methotrexate treatment significantly improve patients' function and health-related quality of life. Significant decrements in functional ability and health-related quality of life were apparent at baseline in this population of methotrexate-naïve patients with active RA, similar to findings reported by Ruta et al and Wells et al (12,13). This is of particular interest given the baseline disease characteristics of the study population. Despite active RA at study entry, this population would be characterized as having mild-to-moderate disease, yet significant improvement in all measures of function and health-related quality of life were evident after 12 months in both active treatment groups when compared with no change in the placebo group.

Leflunomide administration resulted in substantial improvements in the HAQ disability index, MHAQ score, PET top 5 score, and SF-36 physical component score and several scales: physical functioning, bodily pain, general health perceptions, vitality, and social functioning. These changes are clearly important because there was little to no improvement or deterioration in the placebo group. They also reflect significantly greater improvement when compared with methotrexate therapy in this clinical trial, and these improvements correlate, on an individual patient basis, with ACR responses of ≥20% and ≥50%.

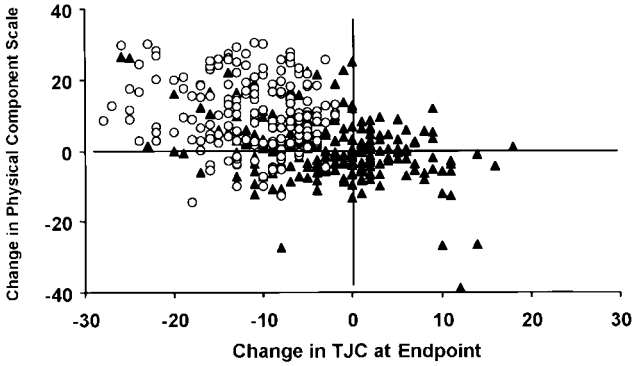

Figure 5. Change in the Short Form 36 (SF-36) physical component score at end point versus American College of Rheumatology (ACR) responders by total joint count (TJC) (leflunomide treatment group, intention-to-treat population). ACR responders (≥20% improvement) are plotted as open circles, and nonresponders as dark triangles. Improvement corresponds to the upper left hand field, indicating corresponding decreases in tender joint count and increases (improvement) in the SF-36 physical component score.

Mean percentage improvements in these measures of function after 12 months of treatment compared favorably with those reported in previous clinical trials of 12–24 weeks' duration and are clinically important. The Outcome Measures in RA Clinical Trials conferences have established an initiative to define "minimum clinically important differences" in specific outcome criteria (23), changes that are apparent and meaningful to both patient and investigator. It has been suggested that improvements of 36% of baseline values, or 18% better than placebo, would be clinically important. Improvements of 22–36% in the HAQ scores (generally, 0.22–0.46 points), ≥0.25 in the MHAQ scores, and 33% in the PET top 5 scores (~5.0 points) are considered clinically important (24,25). In the leflunomide treatment group, all measures met or exceeded these levels: the mean HAQ disability index improved by 0.46 (baseline score of 1.3), reflecting a 35% change, which exceeded placebo by 37%. Mean changes in most of the subscales approached 0.5. Mean MHAQ scores improved by 0.29 (baseline score of 0.78), which was an improvement of 37% and which exceeded the placebo response by 45%. The PET top 5 score improved by 6.9 (baseline of 21.2), reflecting an improvement of 35% or 29% compared with placebo.

Improvements in function/disability measures of this magnitude would be expected to reflect meaningful savings in direct medical costs and cumulative costs of treatment. Fries et al have reported an incremental increase in yearly direct medical costs of $3,000 for each increase of 1.0 in the HAQ disability index (26). Similarly, an increase of 1.0 in the HAQ disability index in the first 2 years of disease resulted in 90% greater disability and 87% more costs over the ensuing 3 years, and 75% greater disability and 74% more costs over the following 8 years (27). Thus, the improvement in function following leflunomide or methotrexate treatment would be expected to result in significant savings to the patient and health system.

Similarly, changes in the SF-36 scales and physical component score appear to reflect clinically meaningful improvement, estimated to be within a 10-point range over 12 months of observation in a longitudinal observational study of patients with systemic lupus erythematosus (28). The results of the current study show that the magnitude of improvement in 6 of the 8 SF-36 scales in the leflunomide group met or exceeded 10 points, with the exceptions of the general health profile, which worsened in the placebo population, and mental health, which reflected numerical, but not statistically significant, improvements in comparison with placebo. Changes in the SF-36 physical component score and individual scales indicate that leflunomide treatment improved health-related quality of life that cannot exclusively be attributed to changes in the performance of physical activities, and reflect that patients had higher energy levels, more vitality, and were able to enjoy more social activities without interference from physical or emotional problems.

Leflunomide administration additionally resulted in significant improvements in work productivity in those patients who engaged in paid employment, housework, and school. These improvements in productivity encompass motivation, commitment, quality, and consistency in the workplace, around the home, and in academic achievements. Reduced productivity is often a major component in the cost of rheumatic disorders such as RA.

Significantly greater improvements in function, performance of activities important to patients, and health-related quality of life were observed in the leflunomide and methotrexate treatment groups when compared with placebo in this 12-month randomized controlled trial. These improvements correlated with the ACR response status and reflected clinically important changes that are meaningful to patients. Improvements in the PET weighted top 5 score reflected significant improvement in performing important physical activities most affected by the underlying RA. Statistically significant changes in the SF-36 physical component score and individual scales indicate that leflunomide and methotrexate treatment improved health-related quality of life.

REFERENCES

1. Allaire SH, Prashker MJ, Meenan RF. The costs of rheumatoid arthritis. PharmacoEconomics 1994;6:513–22.
2. Pincus T, Summey JA, Soraci SA Jr, Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum 1983;26:1346–53.
3. Fries JF, Spitz PW, Young DY. The dimensions of health outcomes: the Health Assessment Questionnaire, disability and pain scales. J Rheumatol 1982;9:789–93.
4. 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.
5. Bombardier C, Buchbinder R, Tugwell P. Efficacy of cyclosporin A in rheumatoid arthritis: long term followup data and the effect on quality of life. Scand J Rheumatol 1992;21:29–33.
6. Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727–35.
7. Bombardier C, Ware J, Russell IJ, Larson M, Chalmers A, Read

JL. Auranofin therapy and quality of life in patients with rheumatoid arthritis: results of a multicenter trial. Am J Med 1986;81: 565–78.
8. Tugwell P, Bombardier C, Buchanan WW, Goldsmith CH, Grace E, Bennett KJ, et al. Methotrexate in rheumatoid arthritis: impact on quality of life assessed by traditional standard-item and individualized patient preference health status questionnaires. Arch Intern Med 1990;150:59–62.
9. Ware JE, Sherbourne CD. The MOS SF-36 item short-form health survey (SF-36): conceptual framework and item selection. Med Care 1992;30:473–83.
10. Tuttleman M, Pillemer S, Tilley B, Fowler S, Buckley L, Alarcon G, et al. Cross sectional assessment of health status instruments in patients with rheumatoid arthritis participating in a clinical trial. J Rheumatol 1997;24:1910–5.
11. Talamo J, Frater A, Gallivan S, Young A. Use of the Short Form 36 for health status measurement in rheumatoid arthritis. Br J Rheumatol 1997;36:463–9.
12. Ruta DA, Hurst JP, Kind P, Hunter M, Stubbings A. Measuring health status in British patients with rheumatoid arthritis. Br J Rheumatol 1998;37:425–36.
13. Wells G, Boers M, Shea B, Tugwell P, Westhovens R, Suarez-Almazor M, et al. Sensitivity to change of generic quality of life instruments in patients with rheumatoid arthritis. J Rheumatol. In press.
14. Silva HT, Morris RE. Leflunomide and the malanonitriloamides. Am J Med Sci 1997;313:289–301.
15. Mladenovic V, Domljan Z, Rozman B, Jajic I, Mihajlovic D, Dordevic J, et al. Safety and effectiveness of leflunomide in the treatment of patients with active rheumatoid arthritis: results of a randomized, placebo-controlled, phase II study. Arthritis Rheum 1995;38:1595–603.
16. Schiff M, Kaine J for the Leflunomide RA Investigators Group, Sharp J, Strand V. X-ray analysis of 12 months treatment of active rheumatoid arthritis with leflunomide compared to placebo or methotrexate [abstract]. Arthritis Rheum 1998;41 Suppl 9:S155.
17. Smolen JS, Kalden JR, Scott DL, and the European Leflunomide Study Group. Efficacy and safety of leflunomide compared with placebo and sulphasalazine in active rheumatoid arthritis: a double-blind, randomized, multicentre trial. Lancet 1999;353: 259–66.
18. Hoechst Marion Roussel, Inc. Leflunomide [package insert]. Kansas City (MO): Hoechst Marion Roussel, Inc; 1998.
19. Strand V, Cohen S, Schiff M, Weaver A, Fleischmann R, Cannon GW, et al, on behalf of the Leflunomide RA Investigators Group. Treatment of active rheumatoid arthritis with leflunomide compared to placebo and methotrexate. Archives Int Med. In press.
20. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
21. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 health status survey manual. Boston: The Health Institute, New England Medical Center; 1993.
22. Tomlin GS, Holm MB, Rogers JC, Kwoh CK. Comparison of standard and alternative health assessment questionnaire scoring procedures for documenting functional outcomes in patients with rheumatoid arthritis. J Rheumatol 1996;23:1524–30.
23. Goldsmith C, Boers M, Bombardier C, Tugwell P, for the OMERACT Committee. Criteria for clinically important changes in outcomes: development, scoring and evaluation of rheumatoid arthritis patient and trial profiles. J Rheumatol 1993;20:561–5.
24. Wells G, Tugwell P, Kraag G, Baker P, Groh J, Redelmeier D. Minimum important difference between patients with rheumatoid arthritis: the patient's perspective. J Rheumatol 1993;20:557–60.
25. Guzman J, Maetzel A, Peloso P, Yeung M, Bombardier C. Disability scores in DMARD trials: what is a clinically important change? [abstract]. Arthritis Rheum 1996;39 Suppl 9:S208.
26. Fries JF, Williams CA, Ramey DR, Singh G. Medical costs are strongly associated with disability levels in rheumatoid arthritis [abstract]. Arthritis Rheum 1995;38 Suppl 9:S187.
27. Singh G, Terry R, Ramey D, Wolfe F, Fries J. Long-term medical costs and outcomes are significantly associated with early changes in disability in rheumatoid arthritis [abstract]. Arthritis Rheum 1996;39 Suppl 9:S318.
28. Thumboo J, Fong KY, Ng TP, Leong KH, Feng PH, Thio ST, et al. The SF-36 is sensitive to changes in quality of life in SLE patients [abstract]. Arthritis Rheum 1998;41 Suppl 9:S225.