Improvement and Longterm Maintenance of Quality of Life During Treatment with Adalimumab in Severe Rheumatoid Arthritis

THOMAS MITTENDORF, BIRGITTA DIETZ, RAIMUND STERZ, HARTMUT KUPPER, MARY A. CIFALDI, and JOHANN-MATTHIAS von der SCHULENBURG

ABSTRACT. Objective. In patients with longstanding severe rheumatoid arthritis (RA) receiving chronic treatment with adalimumab, health related quality of life (HRQOL) was assessed using new instruments [Functional Assessment of Chronic Illness Therapy-Fatigue scale (FACIT-Fatigue) and Health Utilities Index Mark 3 (HUI3)] and a more conventional instrument [Medical Outcomes Study Short Form-36 Health Survey (SF-36)].

Methods. Different measures for collecting patient-reported outcomes were applied simultaneously during the 3-year study period. Sociodemographic and medical history data were assessed at the baseline visit. Clinical examinations (e.g., joint examination and morning stiffness), disease assessments, and HRQOL data were recorded every 8 weeks. For dichotomous and categorical variables, absolute and relative frequencies were calculated. Metric measures were described using mean and standard deviation and/or standard error of the mean. HRQOL data were analyzed using observed cases.

Results. All assessed measures (FACIT-Fatigue, HUI3, SF-36) showed a rapid and statistically significant improvement from baseline following initiation of adalimumab therapy. This effect was maintained over the study period for a mean of 1.6 years in all applied measures. HRQOL data from all tested instruments were significantly correlated with each other.

Conclusion. Chronic therapy with adalimumab improved measures of fatigue and HRQOL in patients with longstanding RA. (First Release Oct 1 2007; J Rheumatol 2007;34:2343–50)

Key Indexing Terms:
RHEUMATOID ARTHRITIS BIOLOGICALS QUALITY OF LIFE COST-EFFECTIVENESS

Rheumatoid arthritis (RA) is an autoimmune disorder that affects approximately 1% of the adult population in industrialized countries1. The disease is characterized by inflammation of multiple joints, resulting in erosive damage of articular cartilage and underlying bone. RA leads to severe functional impairment and a reduction in quality of life. Patients experience reduced physical function, oppressive fatigue, and pain2. Adalimumab (Humira®, Abbott Laboratories, Abbott Park, IL, USA) is the first fully human anti-tumor necrosis factor (TNF) monoclonal antibody available for the treatment of RA. Clinical trials have demonstrated that adalimumab is safe and effective for longterm treatment of patients with RA3–8. Our study focuses on the assessment of health related quality of life (HRQOL) in patients with longstanding, severe RA receiving chronic treatment with adalimumab, using conventional instruments [Health Assessment Questionnaire (HAQ) and the Medical Outcomes Study Short Form-36 Health Survey (SF-36)] and newer instruments [Functional Assessment of Chronic Illness Therapy-Fatigue scale (FACIT-Fatigue) and the Health Utilities Index Mark 3 (HUI3)].

In rheumatology research, HRQOL instruments are commonly used9,10, and substantial data on HRQOL have been collected and analyzed11. Several disease-specific and generic instruments have been validated in rheumatologic diseases, especially in RA12. Four of the more commonly used approaches in measuring HRQOL include (1) measurement of data by disease-specific instruments (e.g., the HAQ as one of the most commonly used functional measures in rheumatology)13; (2) the selection of specific, important components of RA-specific HRQOL (e.g., FACIT-Fatigue)14; (3) generic health profiles (e.g., SF-36)15; and (4) validated health utility instruments (e.g., EuroQol 5D or HUI3)16,17.

In our study, a number of patient-reported outcome (PRO) measures were applied simultaneously in a 3-year, single-study setting. Our aim was to analyze these results from patients with longstanding, severe RA, focusing on measurement of fatigue and health utility. Results describing the HAQ...
from this study were reviewed by van de Putte and colleagues. Cost data also were recorded in the study; these results were presented at the 2006 American College of Rheumatology (ACR) meeting.

MATERIALS AND METHODS

Study design. This long-term, open-label, health outcomes extension study (DE033) included 505 patients with longstanding RA who had received adalimumab therapy during 1 of 6 Phase I-III studies, of which DE026 included a double-blind, randomized, placebo-controlled period of at least 12 weeks in duration (Figure 1). After that, patients received adalimumab 40 mg every other week and were followed for up to 144 weeks. Patients receiving placebo in the preceding DE026 trial were switched to adalimumab treatment at the beginning of this open-label study. The study was performed at 47 investigational sites in Europe (30 sites), Australia (9 sites), and Canada (8 sites), and was conducted in conjunction with clinical study DE018, a multicenter, open-label study evaluating the clinical effectiveness of adalimumab in patients with RA.

The largest of the preceding randomized dose-finding studies, DE026, was a 6-month, Phase III, placebo-controlled study. Because data in DE026 were collected in a manner similar to this study, data from patients in the placebo and adalimumab treatment arms were analyzed as a subgroup (DE026 subgroup) in our study. Outcomes data from studies of adalimumab for the treatment of patients with early RA will be reported separately.

Sociodemographic and medical history data were assessed at the baseline visit. Clinical examination findings (e.g., joint examination, morning stiffness), disease assessments (patients’ and physicians’ global assessments of disease activity and patients’ assessments of pain), and HRQOL data were recorded every 8 weeks. Only validated language translations of all measures were used.

Patients. Patients with RA, as defined by the 1987 revised ACR criteria, were included in the study. Exclusion criteria included (1) pregnant or breastfeeding women; (2) known human immunodeficiency virus-positive status; (3) a history of alcohol or drug abuse within 6 months prior to study entry; (4) ongoing or active clinically relevant infection or any major episode of infection requiring hospitalization or treatment with intravenous antibiotics (within 30 days) or oral antibiotics (within 15 days); and (5) underlying cardiac, pulmonary, metabolic, renal, or gastrointestinal conditions; chronic or latent infectious diseases; immune deficiency; or abnormal laboratory values that, in the opinion of the investigator or the medical monitor, placed the patient at an unacceptable risk.

HRQOL measures. Generic measurement of HRQOL. The SF-36 is the most widely used generic measure of HRQOL. This instrument covers 8 areas of health status, including physical functioning, role–physical, bodily pain, general health, vitality, social functioning, role–emotional, and mental health. The SF-36 scores range between 0 (worst) and 100 (best). In addition, physical component summary (PCS) and mental component summary (MCS) scores can be derived. In RA, minimum clinically important differences (MCID) were defined as a 5- to 10-point change from baseline for the SF-36 subdomains and a 2.5- to 5-point change from baseline for the PCS and MCS score.

Measurement of RA-specific facets of HRQOL. Studies showed that patients regard oppressive fatigue as a major determinant of their overall HRQOL. The FACIT-Fatigue was used to assess fatigue in patients enrolled in our study. The FACIT-Fatigue scale includes 13 specific items linked with fatigue: energy, weakness, lisleness, tiredness, trouble with starting things, trouble with finishing things, energy, activity, sleep, eating, help doing activities, frustration, and social activities. FACIT-Fatigue scores range from 0 to 52, with greater scores representing more fatigue. The instrument has been validated for the general population and for patients with RA. The MCID for FACIT-Fatigue in RA was determined to be at least a 4-point change from baseline.

Measurement of health related utility. The health related utility of patients with RA was evaluated using the HUI3. The first component of the HUI3 is a multi-attribute health status classification system that is used to describe the health status of the patient (e.g., emotion or pain). The second component is a multi-attribute utility function that is used to value the health status as measured within the corresponding multi-attribute health status classification system. These resulting scores in combination with respective timeframes can be converted into quality-adjusted life-years. The score for each level is set by a specific scoring algorithm, which has been based on preferences elicited with special standard-gamble approaches in studies from the general public. Health utility scores are anchored at values from 0 (death) to 1 (perfect health). Scores may reach a value below 0 for disease states that are considered worse than death by the public (i.e., scores can range from −0.36 to 1). The construction of the scale is one of preference or desirability. The more preferable or desirable a health state, the greater its utility. In addition, negative scores are possible and represent health states considered worse than death. Score changes of 0.03 are considered clinically important.

Statistics. All patients enrolled in the study were included in this intention-to-treat (ITT) analysis. For dichotomous and categorical variables, absolute and relative frequencies were calculated. Metric measures are described by mean, standard deviation, and/or standard error of the mean. Spearman rank corre-

![Figure 1](https://www.jrheum.org)
lations were calculated for baseline measures of the Disease Activity Score (DAS28), HAQ, FACIT-Fatigue, HUI3, and SF-36 summary scores (physical and mental). Longitudinal analyses of HRQOL data for the SF-36 were performed by calculating the area under the curve (AUC), obtained by plotting the course of the particular HRQOL measurement. Because the patients remained in the study for different lengths of time, AUC were based on average AUC per visit. For each patient, the AUC for the actual time under observation was determined. Then the average AUC per visit was determined for each patient, i.e., the AUC observed between visits were summed up and divided by the number of visits. Finally, an average AUC per visit across all patients was calculated. Missing HRQOL scores were replaced using the mean imputation technique, where the mean of all available assessments of one patient for the respective measure (except for the baseline value) was used to replace the missing value. In addition, if a death occurred, a replacement of utility values by 0 was performed for the HUI3 analysis. A last-observation-carried-forward approach (LOCF) was used in some analyses for the subgroup of patients from study DE026 who participated in the placebo-controlled study with 40 mg adalimumab every other week for 26 weeks (n = 99). Differences between placebo and adalimumab groups during the randomized controlled portion of the DE026 study were compared using t-tests after confirming normal distribution of the data. The Mann-Whitney U-test was used for between-group comparisons of non-normally distributed data. All statistical analyses were performed using SAS® version 8.2 (SAS Institute Inc., Cary, NC, USA).

RESULTS

Patients. A total of 505 patients were enrolled in our study, with the greatest percentage of patients enrolled in Germany (n = 153). Patients were recruited from several preceding dose-finding studies; most patients originated from DE026 (Figure 1).

On average, patients participated in the health outcomes study for 1.6 years (mean 1.57 ± 0.63 yrs; median 1.8 yrs). The number of completed questionnaires varied between instruments. For example, for the HUI3, 321, 406, 369, and 371 completed questionnaires were collected after 72, 96, 120, and 144 weeks, respectively. Baseline patient characteristics are provided for 2 patient populations (Table 1): the DE026 subgroup and the overall DE033 population. Baseline data for the DE026 subgroup (n = 99) include data from patients who were naive to biologic treatment and who received the same adalimumab dosage regimen (adalimumab 40 mg every other wk) as the overall population for 26 weeks. Data from the DE026 cohort are included in the entire study population (n = 505). There were no statistically significant differences between placebo and adalimumab groups at baseline from the DE026 study for all measures.

Overall, most of the patients in the entire study group had longstanding, severe RA, with a mean duration of 12.4 years and an average of about 4 failed previous disease modifying antirheumatic drugs. About 75% of patients were female, confirming a similar sex distribution of patients with RA among industrialized countries.

Short Form-36 Health Survey. Among the DE026 subgroup, there were no significant differences in SF-36 scores between the placebo and adalimumab treatment groups at baseline. At Week 26, patients receiving adalimumab achieved significant improvement in all SF-36 subdomains; changes were statistically significant compared with placebo and compared with baseline (Figure 2). Increases in all SF-36 subdomains in the DE026 subgroup were maintained for more than 3 years (Figure 3). Table 2 provides a comparison of the baseline SF-36 scores for patients in the DE026 study (reflecting scores prior to adalimumab treatment), baseline SF-36 score of the entire study population, and SF-36 scores after 144 weeks of adalimumab treatment. SF-36 scores from the DE026 subgroup (Figure 3) were consistent with SF-36 scores from the entire study population (Table 2). All increases in the SF-36 subdomains of the SF-36 were clinically relevant, meaning the MCID was achieved.

Table 1. Baseline demographic data from DE026 subgroup and overall DE033 population.

<table>
<thead>
<tr>
<th>Variable</th>
<th>DE026 Subgroup Receiving 40 mg Adalimumab for 26 Weeks Prior to DE033 (n = 99)</th>
<th>Overall DE033 Population (n = 505)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female, n (%)</td>
<td>79 (79.8)</td>
<td>390 (77.2)</td>
</tr>
<tr>
<td>Age (yrs), median (range)</td>
<td>54 (19–80)</td>
<td>55 (22–82)</td>
</tr>
<tr>
<td>Weight (kg), mean ± SD</td>
<td>69.7 ± 13.92</td>
<td>69.12 ± 12.90</td>
</tr>
<tr>
<td>Height (cm), mean ± SD</td>
<td>165.9 ± 8.47</td>
<td>166.06 ± 8.90</td>
</tr>
<tr>
<td>Caucasian, n (%)</td>
<td>95 (96.0)</td>
<td>498 (98.6)</td>
</tr>
<tr>
<td>BMI (kg/m²), mean ± SD</td>
<td>25.4 ± 4.96</td>
<td>25.06 ± 4.34</td>
</tr>
<tr>
<td>Employed, n (%)</td>
<td>NA</td>
<td>152 (30.1)</td>
</tr>
<tr>
<td>Retired, n (%)</td>
<td>NA</td>
<td>220 (43.6)</td>
</tr>
<tr>
<td>RA duration, yrs, mean ± SD</td>
<td>10.3 ± 7.05</td>
<td>12.36 ± 7.69</td>
</tr>
<tr>
<td>No. previous DMARD failed, mean ± SD</td>
<td>3.8 ± 1.77</td>
<td>3.7 ± 1.82</td>
</tr>
<tr>
<td>TJC (0–68), mean ± SD</td>
<td>33.81 ± 15.97</td>
<td>14.58 ± 14.89</td>
</tr>
<tr>
<td>SJC (0–66), mean ± SD</td>
<td>21.02 ± 10.97</td>
<td>8.23 ± 8.18</td>
</tr>
<tr>
<td>DAS28, mean ± SD</td>
<td>5.39 ± 1.62</td>
<td>4.58 ± 1.58</td>
</tr>
</tbody>
</table>

* Eight patients missing. † Two patients missing. ‡ Including self-employed. § Baseline data for both cohorts differ because patients in the overall population had already been treated with adalimumab. BMI: body mass index; DMARD: disease modifying antirheumatic drugs; DAS28: Disease Activity Score 28; RA: rheumatoid arthritis; SJC: swollen joint count; TJC: tender joint count.
**Figure 2.** SF-36 health profile scores at baseline and after 26 weeks of treatment with adalimumab or placebo in the DE026 subgroup (n = 99). *Adalimumab p < 0.01 vs baseline and vs placebo. †Adalimumab p < 0.05 vs baseline and vs placebo. Placebo-treated patients did not achieve statistical significance vs baseline. Last observation carried forward. ‡US population norms from Ware, et al. Every other week.

**Figure 3.** SF-36 health profile scores in the DE026 subgroup over 3 years (n = 99). *Adalimumab p < 0.02 vs baseline for all domains except role-physical, which was p < 0.05; on and after Week 26. Last observation carried forward.
Figure 4. Change from baseline in FACIT-Fatigue scores in the DE026 subgroup over 3 years (n = 99). *p < 0.001 vs baseline. †p < 0.01 vs placebo. Last observation carried forward. MCID: minimum clinically important difference; SEM: standard error of the mean.

Figure 5. Change from baseline in HUI3 scores in the DE026 subgroup over 3 years (n = 99). *p < 0.001 vs baseline. †p < 0.05 vs placebo. Last observation carried forward. MCID: minimum clinically important difference; SEM: standard error of the mean.
Correlation analysis. All assessed instruments were significantly correlated with each other at baseline. As expected, both measures of physical function (SF-36 physical and HAQ) were highly correlated (Spearman rank correlation 0.733; p < 0.0001). The FACIT-Fatigue was highly correlated with the HUI3 (0.671; p < 0.0001). The SF-36 mental was only slightly correlated with physical function (HAQ and PCS) or clinical response (DAS28), and moderately correlated with FACIT-Fatigue (Spearman rank correlation −0.272, p < 0.0001; 0.162, p < 0.001; −0.154, p < 0.01; and 0.513, p < 0.0001, respectively) (Table 3).

DISCUSSION

This health outcomes trial was conducted as a companion study to the adalimumab clinical trials in patients with longstanding RA. HRQOL was assessed using specific quality of life questionnaires (SF-36, FACIT-Fatigue, HUI3). All HRQOL measures reflected a rapid and statistically significant improvement from baseline in HRQOL after initiation of adalimumab therapy. This finding was consistent with the results reported previously for the disease-specific measure (HAQ)18.

Improvements in HRQOL measures were considered clinically meaningful and were maintained for, on average, 1.6 years. Mean SF-36, FACIT-Fatigue, and HUI3 scores improved rapidly and remained stable during the entire treatment period. SF-36 scores indicated that patients with late-stage RA are especially impaired in physical functioning, physical role, bodily pain, general health, and vitality. This finding is consistent with previous findings for the SF-3623,24 and for the FACIT-Fatigue and HUI325,26. The maintenance of the utility values over time is important to clinical practice because disability and clinical measures, such as joint and bone destruction, progress over time, especially among patients with longstanding RA, such as those enrolled in our study27.
Our study provided the only information to date on the positive and clinically meaningful effects of longterm treatment with adalimumab or any other TNF antagonist on certain facets of HRQOL. This trial was the first to measure longterm effects of a TNF antagonist on fatigue. Measuring fatigue as a marker of impairment proved to be valid: FACIT-Fatigue results correlated significantly with other, better-established HRQOL measures. To date, the HUI3 previously has been used only once in a clinical trial of patients with RA receiving TNF-antagonist treatment. Consistent with the results of our study, adalimumab provided significant improvement in HUI3 in patients with longstanding RA during the 12-month duration of the study. Our study demonstrated that with the instruments used in our study, even relatively small improvements can be observed. Other studies that did not explicitly evaluate TNF-antagonist therapy previously showed the sensitivity to small improvements: What is the relative importance of different HRQOL facets describes an area in which patients with RA exhibit significant impairment compared with the general population. The FACIT-Fatigue also provides researchers with insight into an issue of particular concern to patients: oppressive fatigue. The Outcome Measures in Rheumatology [Clinical Trials] (OMERACT) group will research this issue further in the coming years.

The strengths of our study are the long duration, the cohort size, the parallel assessment of multiple PRO measures in identical timeframes, the multinational approach in industrialized countries (all with existing high treatment standards for patients with RA), and the combination of PRO results within clinical trials, this study supports the use of at least 2 additional measures — the SF-36 and the FACIT-Fatigue — to further characterize the disease burden. The SF-36 was validated for use in RA and can be used to measure health utility, and reflects 8 different facets of HRQOL. Each of these facets describes an area in which patients with RA exhibit significant impairment compared with the general population. The FACIT-Fatigue also provides researchers with insight into an issue of particular concern to patients: oppressive fatigue. The Outcome Measures in Rheumatology [Clinical Trials] (OMERACT) group will research this issue further in the coming years.

The results from this longterm observational study may be used as input data for health economic models evaluating the longterm effects of biologic treatment. Recent studies have shown favorable cost-effectiveness ratios for biologic treatments. These studies could be further improved by incorporating longterm HRQOL outcomes. Valuable modeling questions may be explored, such as the impact of adalimumab on future direct costs, indirect costs to society, progressive increases in costs, and the future costs of newly diagnosed patients. Further, lifetime economic and PRO models could be established using recent data demonstrating reduced mortality during TNF-antagonist therapy.

Based on the results of our study, new questions arise that await to be addressed in the future with results from current trials: What is the relative importance of different HRQOL measures? Is there a HRQOL measure that is superior to other measures? Is there a HRQOL measure that is superior to other measures?

### Table 3. Baseline correlations (Spearman rank correlation coefficients; n = 505).

<table>
<thead>
<tr>
<th>FACIT-Fatigue</th>
<th>HAQ</th>
<th>HU13</th>
<th>SF-36 Mental</th>
<th>SF-36 Physical</th>
<th>SF-36 Pain</th>
<th>DAS28</th>
</tr>
</thead>
<tbody>
<tr>
<td>FACTIT-Fatigue 1.00000</td>
<td>-0.44244</td>
<td>0.67075</td>
<td>0.51296</td>
<td>0.42704</td>
<td>0.45558</td>
<td>-0.42028</td>
</tr>
<tr>
<td>HAQ &lt; 0.0001</td>
<td>1.00000</td>
<td>-0.58887</td>
<td>-0.27191</td>
<td>-0.73307</td>
<td>-0.65008</td>
<td>0.34137</td>
</tr>
<tr>
<td>HU13 &lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>SF-36 mental &lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>0.30104</td>
<td>0.48078</td>
<td>0.48762</td>
<td>-0.49559</td>
</tr>
<tr>
<td>SF-36 physical &lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>0.16172</td>
<td>0.42441</td>
<td>-0.15394</td>
</tr>
<tr>
<td>SF-36 pain &lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>DAS28 &lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

DAS28: Disease Activity Score 28; FACIT-Fatigue: Functional Assessment of Chronic Illness Therapy-Fatigue scale; HAQ: Health Assessment Questionnaire; HU13: Health Utilities Index Mark 3; SF-36: Short-Form 36 Health Survey.
instruments in terms of predicting future disability, fatigue, or costs? Which facet of HRQOL should be measured? Are values maintained after initial treatment? What is the role of the patient with RA in the future? Should all patient-relevant outcomes be measured? What are the most important outcomes for chronically ill patients: clinical outcome or PRO?

ACKNOWLEDGMENT

We acknowledge the editorial support of Lori Lush, PharmD, and Michelle Metelo, PharmD, of JK Associates, Inc., in the preparation of this article.

REFERENCES
