Value of the time trade off method for measuring utilities in patients with rheumatoid arthritis


Abstract

Objective—To assess the feasibility, reliability, and validity of the time trade off (TTO) in patients with rheumatoid arthritis (RA).

Methods—The TTO was applied in 194 patients with RA with increasing difficulty in performing activities of daily living. The test-retest reliability was determined in 35 of these patients and was calculated by the intraclass correlation coefficient (ICC). Construct validity was evaluated with the following sets of variables: measures of utility (rating scale), quality of life (RAND 36 item Health Status Survey (RAND-36) and RAQoL), functional status (Health Assessment Questionnaire, grip strength, and walk test), and disease activity (doctor’s global assessment, disease activity score, pain, and morning stiffness).

Results—Ten patients (5%) did not complete the TTO. The median value of the TTO was 0.77 (range 0.03–1.0). The test-retest ICC of the TTO was 0.85 (p<0.001). Construct validity testing of the TTO showed poor to moderate correlations (Spearman’s r, between 0.19 and 0.36, p<0.01) with all outcome measures except for the subscale role limitation (physical problem) of the RAND-36, the walk test, the doctor’s global assessment of disease activity, and morning stiffness. Multiple regression analysis showed that only 17% of the variance of the TTO scores could be explained.

Conclusions—The TTO method appeared to be feasible and reliable in patients with RA. The poor to moderate correlations of the TTO with measures of quality of life, functional ability, and disease activity suggest that the TTO considers additional attributes of health status. This may have implications for the application of the TTO in clinical trials in patients with RA. 

Rheumatoid arthritis (RA) is a chronic disease, which is associated with considerable physical impairment, work disability, and frequent healthcare utilisation. The continuing development of new treatment strategies for patients with RA and the increasing scarcity of resources make it necessary to prove that treatment strategies are not only effective but also that they are cost effective. Therefore, a growing number of clinical studies in RA include economic evaluations. One method of economic evaluation is the cost-utility analysis. In this method a utility is used as a global, health related, quality of life measure. A utility is defined as the level of subjective satisfaction, distress, or desirability that people associate with a particular outcome. It is expressed as a single value on a scale ranging from 0 (usually perfect health) to 1 (usually perfect health). The values that subjects attach to these health states are called the utilities of health states. They are used as quality adjustment factors. Until now, there has been little information about the meaning of utility values in relation to more common outcome measures in patients with RA.

The three most widely used methods of utility measurement are the standard gamble (SG), the time trade off (TTO), and the rating scale (RS) or visual analogue scale (VAS). With the SG, the respondent is asked to make a choice between two options. The first option is the certainty of living for the rest of one’s life in a particular health condition. The other option is a gamble with two possible outcomes, living for the rest of one’s life in perfect health or immediate death. The chances in the gamble are varied to determine the point at which a respondent is indifferent about the choice between the certain option and the gamble.

The TTO asks the respondent to value health states in terms of duration of life in a state of perfect health that would be equivalent to some period in a particular health condition—for example, the patient’s own health state. In contrast with the SG, in the TTO no element of uncertainty is involved. The RS does not reflect any trade off that a patient may be willing to make in order to obtain better health, either in terms of risk or in life years. Utilities derived by the TTO and SG are more closely linked to the theoretical foundations of utility measurement than is the case with the RS.

However, it was found that the SG is less responsive in patients with chronic musculoskeletal pain. Moreover, the TTO has the advantage of being easier to understand and taking less time to administer than the SG.

The TTO has been used in some studies of patients with rheumatic diseases. However, there are no studies in which the feasibility, reliability, and validity of this instrument have been described in patients with RA. Therefore, the aim of this study was to assess these properties in a group of patients with RA.

Patients and methods

Patients

Data were gathered as part of a randomised controlled trial, in which the cost effectiveness
of three types of care (inpatient multidisciplinary care, multidisciplinary day care, and outpatient care coordinated by a rheumatologist and a nurse practitioner) for patients with RA was compared. The patients were recruited between December 1996 and January 1999 in the Leiden University Medical Centre and five non-academic hospitals within the region of Leiden. Two hundred and ten consecutive patients who visited the outpatient clinic of the rheumatology departments of these hospitals, and who satisfied the inclusion criteria, were included. The inclusion criteria were RA as defined by the 1987 revised ARA criteria and increasing difficulty in performing activities of daily living for at least six weeks. Functional problems were not solvable by the rheumatologist alone or by a single health professional. Patients who had medical complications of RA requiring immediate admission to hospital and patients who were unable to reach the hospital before 10 am were excluded. The medical ethics committees of all participating hospitals approved the study protocol. All patients gave written informed consent.

**METHODS**

Two trial nurses who were blinded to the patient’s treatment status carried out all assessments. Assessments were made at baseline and weeks 6, 12, 26, 52, 78, and 104. For the utility study, only the data gathered at baseline were used.

**Sociodemographic characteristics and clinical data**

The following data were recorded at baseline: sex, age, disease duration, rheumatoid factor, status of living (living alone or with partner), and level of education (categorised as low: up to and including lower technical and vocational training; medium: up to and including secondary technical and vocational training; high: up to and including higher technical and vocational training and university).

**Utility measures**

Two measures of utility were used—the TTO and the RS. The TTO is a method in which the patients were asked how many years (x) in perfect health they would consider equivalent to their remaining life expectancy (y) in the current health state. The number of years in perfect health is varied and the point of indifference between his own health state for the remaining life expectancy and perfect health for a shorter period of time was used to calculate the utility score. The utility was then calculated as x/y. A standardised interview format was used, in which the patient was shown a card with a description of perfect health (physical: no complaints or difficulties in performing activities of daily living; psychological: always good and positive feelings and mood; social: all sort of social activities can be performed and contacts with other people are good). To determine the point of indifference between his own health state for the remaining life expectancy and perfect health for a shorter period of time, a visual aid was used. To determine the life expectancy of a patient, the age-specific life expectancy from the general population was used (Netherlands Central Bureau for Statistics, 1995). If it was not possible to complete the method, the patients were asked for the reason. To investigate the test-retest reliability, in 35 consecutive patients who were randomly allocated to receive the inpatient or day care, the TTO method was applied again two weeks after baseline assessments. Both the baseline and the retest assessment took place just before the start of the trial intervention and therefore it was expected that the patients’ condition was stable between these two times.

The RS is a method in which the patients are asked to place their current health state alongside a 0–100 mm vertical line anchored by the “worst imaginable health state” (bottom) and perfect health (top).

**Quality of life measures**

Two measures of quality of life (QoL) were used, the RAND-36 and the RAQoL. The RAND-36 item Health Survey 1.0 (RAND-36) contains subscales for physical functioning, social functioning, role limitations (physical problem), role limitations (emotional problem), mental health, vitality, pain, and general health perception. Each scale generates a score from 0 to 100, with a higher score indicating better health. The RAND-36 may be converted to two summary scales: the physical and mental component summary scales.13 The RAND-36 includes the same items as those in the MOS-SF 36 and although the scoring procedures are somewhat different, the effects of these on the scores are minimal.13 The RA-specific QoL instrument (RAQoL), which consists of 30 items with a yes/no (1/0) response format, measures different areas of life, including moods and emotions, social life, hobbies, every day tasks, personal and social relationships, and physical contact.17 The overall score is the sum of the scores of individual items, with a lower score indicating better QoL. The RAQoL was completed by the patient him or herself. The RAND-36 and the RAQoL have been validated for use in the Netherlands.17 18

**Functional measures**

Functional ability was measured by (a) the Dutch version of the Health Assessment Questionnaire (HAQ),19 (b) the grip strength (mean grip strength of the right and the left hand) (Accoson vigorimeter), and (c) the walk test (time needed to walk 50 feet (15 m)).20

**Measures of disease activity**

Measures of disease activity included (a) the patient’s global assessment of disease activity, pain, and morning stiffness, all three measured on a horizontal visual analogue scale (VAS, range 0–10 cm). The left anchors were no disease activity, no pain, and no stiffness, respectively, and the right, worst imaginable disease activity, severe pain, and severe stiffness, respectively. (b) The doctor’s estimation of disease activity, measured on a VAS (range 0–10 cm; anchored on the left by no disease
Table 1 Characteristics of 194 patients with rheumatoid arthritis

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No of patients (%)</th>
<th>Median (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>194 (100)</td>
<td>58 (22–85)</td>
</tr>
<tr>
<td>Duration of disease (years)</td>
<td>194 (100)</td>
<td>1.84 (0.02–46.55)</td>
</tr>
<tr>
<td>Living alone</td>
<td>194 (100)</td>
<td>53 (27)</td>
</tr>
<tr>
<td>Education level*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>104 (54)</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>65 (33)</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>25 (13)</td>
<td></td>
</tr>
</tbody>
</table>

*Low = up to and including lower technical and vocational training; medium = up to and including secondary technical and vocational training; high = up to and including higher technical and vocational training and university.

Table 2 Median values and interquartile ranges of measures of utilities, quality of life, functional status, and disease activity in 194 patients with rheumatoid arthritis

<table>
<thead>
<tr>
<th>Measure</th>
<th>Median</th>
<th>Interquartile range</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTO† (0–1)*</td>
<td>0.77</td>
<td>0.65–0.91</td>
</tr>
<tr>
<td>Rating scale (0–100)/100</td>
<td>0.60</td>
<td>0.45–0.70</td>
</tr>
<tr>
<td>Quality of life</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND-36 (0–100):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical functioning</td>
<td>35</td>
<td>20–60</td>
</tr>
<tr>
<td>Social functioning</td>
<td>64</td>
<td>38–81</td>
</tr>
<tr>
<td>Role limitations physical functioning</td>
<td>0</td>
<td>0–25</td>
</tr>
<tr>
<td>Role limitations emotional problems</td>
<td>33</td>
<td>0–100</td>
</tr>
<tr>
<td>Mental health</td>
<td>64</td>
<td>52–76</td>
</tr>
<tr>
<td>Vitality</td>
<td>45</td>
<td>35–55</td>
</tr>
<tr>
<td>Pain</td>
<td>45</td>
<td>22–56</td>
</tr>
<tr>
<td>General health perception</td>
<td>45</td>
<td>35–50</td>
</tr>
<tr>
<td>Summary scale mental</td>
<td>57</td>
<td>35–80</td>
</tr>
<tr>
<td>Summary scale physical</td>
<td>30</td>
<td>18–41</td>
</tr>
<tr>
<td>RAQoL (0–30)</td>
<td>16</td>
<td>11–21</td>
</tr>
<tr>
<td>Functional status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAQ† (0–3)</td>
<td>1.25</td>
<td>0.88–1.75</td>
</tr>
<tr>
<td>Grip strength (mm Hg)</td>
<td>150</td>
<td>111–217</td>
</tr>
<tr>
<td>Walk test (s)</td>
<td>11</td>
<td>9–14</td>
</tr>
<tr>
<td>Disease activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disease activity VAS†(investigator)</td>
<td>38</td>
<td>24–52</td>
</tr>
<tr>
<td>Pain VAS</td>
<td>54</td>
<td>33–70</td>
</tr>
<tr>
<td>Morning stiffness VAS</td>
<td>55</td>
<td>30–74</td>
</tr>
<tr>
<td>DAS†</td>
<td>5.6</td>
<td>4.7–6.4</td>
</tr>
</tbody>
</table>

*Measured in 184 patients, for all other variables <1.5% missing data.
†TTO = time trade off; HAQ = Health Assessment Questionnaire; VAS = visual analogue scale (range 0–100), DAS = disease activity score.

To investigate a possible influence of patient characteristics, such as age, sex, status of living, and level of education, on the TTO scores the associations between the TTO and these characteristics were quantified by Spearman’s correlation coefficients.

In the general population it has been found that the relation between measures obtained by the RS method and those obtained by the TTO method can be described by the power function RS = 1 – (1 – TTO)^0.82. As support for the validity of the TTO method, it was further assessed whether after this transformation the RS scores were similar to the TTO scores. The differences between TTO scores and the transformed RS scores were tested for statistical significance by the Wilcoxon signed ranks test.

Construct validity of the TTO was evaluated by quantifying Spearman’s correlation coefficients between the TTO method and measures of utility, QoL, functional status, and disease activity. With 194 observations the power to detect correlation coefficients of 0.2 or higher is 0.80.

To illustrate the ability of the TTO to discriminate between worse and better health state outcomes, the following measures were divided into two categories (<median and >median values): the RAND-36 (physical and mental), the RAQoL, the VAS pain, and the DAS scores. The HAQ was divided into four parts according to quartiles. Differences of TTO scores between the two health state categories were tested for statistical significance by the Mann-Whitney U test. The differences between the median values of the TTO scores in the four HAQ categories were analysed by analysis of variance.

To determine the contribution of several sets of variables to the explained variance of the TTO, a multiple regression analysis was performed with the following sets of variables: patient characteristics, measures of QoL, functional status, and disease activity. Firstly, we studied sets of variables separately to reduce the problem of multiple testing and to estimate the roles of different types of variables (method enter). The dependent variable was the TTO. Secondly, within each set of variables, the contribution of individual variables was calculated (stepwise forward). Subsequently, the variables from the different sets of variables that contributed significantly to the explained variance (p<0.001) were again analysed into one final model (stepwise forward).

Results

PATIENT CHARACTERISTICS

Two hundred and ten patients were included in the randomised clinical trial. For logistic reasons, the TTO method was only used after the 16th patient, and was applied in all the following 194 patients. Table 1 shows the sociodemographic and clinical characteristics of the patients. Table 2 shows the median values of measures of utilities, QoL, functional status, and disease activity.
Table 3 Spearman’s correlation coefficients between the time trade off (TTO) and health status outcomes

<table>
<thead>
<tr>
<th>Utility</th>
<th>TTO score (0-1)</th>
<th>TTO (0-100)</th>
<th>Quality of life</th>
<th>Disease activity</th>
<th>Pain management</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAND-36 (0-100)</td>
<td>0.18**</td>
<td>0.25**</td>
<td>Physical functioning</td>
<td>0.26**</td>
<td></td>
</tr>
<tr>
<td>Social functioning</td>
<td>0.25**</td>
<td>0.26**</td>
<td>Role limitations physical functioning</td>
<td>0.26**</td>
<td></td>
</tr>
<tr>
<td>Mental health</td>
<td>0.27**</td>
<td>0.26**</td>
<td>Vitality</td>
<td>0.26**</td>
<td></td>
</tr>
<tr>
<td>General health perception</td>
<td>0.18</td>
<td>0.26**</td>
<td>Pain</td>
<td>0.26**</td>
<td></td>
</tr>
<tr>
<td>Summary scale physical</td>
<td>0.26**</td>
<td>0.26**</td>
<td>Summary scale physical</td>
<td>0.26**</td>
<td></td>
</tr>
<tr>
<td>HAQ†</td>
<td>−0.34**</td>
<td>−0.34**</td>
<td>Functional status</td>
<td>−0.29**</td>
<td></td>
</tr>
<tr>
<td>Grip strength</td>
<td>0.26**</td>
<td>0.26**</td>
<td>Walk test</td>
<td>−0.07</td>
<td></td>
</tr>
</tbody>
</table>

FEASIBILITY
Ten of the 194 patients (5%) did not complete the TTO method. Four patients thought the method was too confronting, three patients could not imagine a situation of perfect health, and three patients mentioned religious reasons. There were no significant differences in age, disease duration, education level, and sex between those who completed and those who did not complete the TTO method. However, there was a tendency for higher age (median 65 years) and longer duration of disease (median 2.6 years) in the group of patients who did not complete the TTO than in the group who did complete it (median age 58 years and duration of disease 1.7 years, respectively). Completion of the TTO method took less than 10 minutes.

RELIABILITY
The two week test-retest reliability was assessed in 35 patients. The ICC for the comparison of both TTO scores was 0.85 (p<0.001).

VALIDITY
Association between utilities and patient characteristics
No statistically significant correlations were found between the TTO and the patient characteristics (age, sex, status of living, level of education) (data not shown).

Transformation of RS into TTO
The scores derived by the TTO method were significantly higher than those derived by the RS method (p<0.001). The median value of the transformed RS scores, using the power function RS = 1 − (1 − TTO)^0.62, was 0.77 (range 0.3–1.00). There was no significant difference between the median of transformed RS scores and the median of actual scores of the TTO (median 0.77, range 0.3–1.00) (p=0.384).

Association between utilities and health status outcome measures
Poor to moderate correlations were found between the TTO scores and the RS, and measures of QoL, functional status, and disease activity (Table 3). All correlations between the TTO and the different outcome measures indicated that a better outcome for utilities, QoL, functional status, and disease activity was associated with higher TTO scores. Tables 4 and 5 illustrate the discriminating properties of the TTO for two (“worse and better”) categories of the RAND-36, RAQoL, VAS pain, and DAS scores and for four categories of the HAQ, respectively.

Regression analysis
In a multiple regression analysis (method enter) the set of variables of QoL, functional status, and disease activity explained 19%, 10%, and 10% of the variance of the TTO, respectively. After stepwise forward analysis within each set of variables it appeared that the following variables contributed significantly to the explained variance (p<0.001): the RAQoL and the subscale vitality of the RAND-36 (set QoL), the HAQ score (set functional status), and the VAS pain (set disease activity). Stepwise forward analysis of these four variables into one model resulted in a significant contribution of 17% to the explained variance of the TTO of two variables: the RAQoL and the subscale vitality of the RAND-36 (p<0.001).

Discussion
The purpose of this study was to assess whether the TTO is a feasible, reliable, and valid instrument to measure utilities in patients...
with RA. The feasibility of applying the TTO method was generally satisfactory. Only 10 (5%) patients did not complete the TTO method. These patients thought the method was too confronting, or they could not imagine a health status of perfect health, or they would not answer for reasons of a religious nature. The possibility of difficulties of religious nature has been put forward by other authors. There was a tendency for older patients with longer disease duration to have difficulty completing the TTO method in comparison with those who did complete it. This age phenomenon has also been found in other patient groups in which the TTO and the SG were applied. The reliability of the TTO method was good. In other studies of utilities in patients with rheumatic diseases, comparable values of the ICC have been found for the TTO and for other measures of utility. The median value of 0.77 for the TTO method in this group of patients corresponds with values derived by the same method in other groups of patients with RA, osteoarthritis, and SLE. A significant difference was found between scores derived by the TTO method and by the RS. This is consistent with the well known empirical finding that TTO scores are generally higher than RS scores. 27–29 The difference may be explained by the difference in tasks. The TTO requires a trade off, whereas with the RS no consequences about life are linked for the respondent. Therefore, the patients who rate their own health status on a 0–100 mm line may use this line more equally than is the case with the TTO method. 27–29

On the basis of Spearman’s correlation coefficient between the TTO and the RS, we can conclude that for individual subjects the RS method is poorly correlated with the TTO method. However, for groups, by using the power function RS = (1 + (1 – TTO))0.62, the median values of the TTO and the transformed RS scores did not significantly differ from each other. This is a finding previously described in other patient groups and supports the validity of the TTO in this group of patients with rheumatic diseases.

Evaluation of the construct validity showed that the TTO is only poorly to moderately related to measures of QoL, functional ability, and disease activity. The correlation coefficients between the TTO and the RAND-36 corresponded with results found in studies with other patient groups in which the MOS-SF 36 and the RAND-36 were used. 30–31 Regression analysis shows that the combination of the subscale vitality of the RAND-36 and the RAQoL explained only a small part of the variance of the TTO method. The results of the RAND-36 corresponded with those of other studies. 32–34 The results of the correlations and regression analysis suggest that the TTO is influenced by other factors that may or may not be related to health status. Factors that may possibly contribute to utilities are adaptation to illness, knowledge about the disease, previous experience with disease related problems, and individual beliefs about health in general. Moreover, factors such as family life (partner, children), social networks, and financial status may have a role. These factors are probably difficult to influence with current interventions.

In conclusion, the TTO method appeared feasible and reliable in a group of patients with RA with increasing difficulty in activities of daily living. Although related to measures of QoL, functional ability, and disease activity it seems that the TTO utilities deal with additional attributes of health status. This may have implications for the application of the TTO in clinical trials, because the sensitivity of the method for detecting clinically relevant differences may be limited. Further research into the usefulness of the method in different groups of patients with RA and in RA clinical trials, including cost-utility analysis, is warranted.

We thank the following rheumatologists who participated in the study: JM de Jonge-Bok, Geerose Hert Ziekenhuis, Leusda; Dr T. Malieé, Kennemer Gasthuis, Haarlem; JF Terwel, Spaarne Ziekenhuis, Haarlem; Dr AJ Feeters and Dr D van Zeven, Reumer van Graaf Gasthuis, Delft; JAM Eswals, Rosendaal; J. B. M. J. Ziekenhuis, Den Haag; GA Kok and Dr HK Rondal, Leyenburg Ziekenhuis, Den Haag; I Speyer and Dr ML Westdij, Bronovo Ziekenhuis, Den Haag. In addition, we are indebted to all rheumatology nurse specialists from the participating hospitals. We would like to acknowledge I Henkes and B Oud for making the assessments and M de Best-Bosz, H Ravensbergen-van de Berg, and A van der Blom for secretarial assistance.

This study was financially supported by Het Nationaal Reuma Fonds, The Dutch Arthritis Association, Grant No 931.


4 Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv Res 1972;7:118–33.


Value of the time trade off method for measuring utilities in patients with rheumatoid arthritis


Ann Rheum Dis 2000 59: 892-897
doi: 10.1136/ard.59.11.892

Updated information and services can be found at:
http://ard.bmj.com/content/59/11/892

These include:

References
This article cites 26 articles, 3 of which you can access for free at:
http://ard.bmj.com/content/59/11/892#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Connective tissue disease (4253)
- Degenerative joint disease (4641)
- Immunology (including allergy) (5144)
- Musculoskeletal syndromes (4951)
- Rheumatoid arthritis (3258)
- Pain (neurology) (883)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/