Unstable Preferences:

A Shift in Valuation or an Effect of the Elicitation Procedure?

SYLVIA J. T. JANSEN, MA, ANNE M. STIGGELBOUT, PhD, PETER P. WAKKER, PhD, MARIANNE A. NOOIJ, MD, EVERT M. NOORDIJK, MD, PhD, JOB KIEVIT, MD, PhD

Objective. Many studies suggest that impaired health states are valued more positively when experienced than when hypothetical. This study investigated to what extent this discrepancy occurs and examined four possible explanations: non-corresponding description of the hypothetical health state, new understanding due to experience with the health state, valuation shift due to a new status quo, and instability of preference. *Patients and methods.* Fifty-five breast cancer patients evaluated their actually experienced health state, a radiotherapy scenario, and a chemotherapy control scenario before, during, and after postoperative radiotherapy. Utilities were elicited by means of a visual analog scale (VAS), a chained time tradeoff (TTO), and a chained standard gamble (SG). *Results.* The discrepancy was found for all methods and was statistically significant for the TTO (predicted utilities: 0.89, actual utilities: 0.92, $p \leq 0.05$). During radiotherapy, significant differences ($p \leq 0.01$) were found between the utilities for the radiotherapy scenario and the actual health state by means of the VAS and the SG, suggesting non-corresponding description as an explanation. The utilities of the radiotherapy scenario and the chemotherapy control scenario remained stable over time, and thus new understanding, valuation shift, and instability could be ruled out as explanations. *Conclusion.* Utilities obtained through hypothetical scenarios may not be valid predictors of the value judgments of actually experienced health states. The discrepancy in this study seems to have been due to differences between the situations in question (non-corresponding descriptions). *Key words:* stability; utility assessment; standard gamble; time tradeoff; breast cancer; chemotherapy; radiotherapy. **(Med Decis Making 2000;20:62–71)**

Utility assessment is an integral part of cost–utility analyses and decision analyses. A central issue in utility measurement is whether utilities are stable, and thus are the same when measured at different time points. If utilities are stable, then measurements taken at any specific time point can be used to assess quality-adjusted life expectancy. If utilities are not stable, then the question arises which utilities are appropriate: those of patients imagining an outcome, those of patients who have experienced the outcome in the past, or those of patients who are actually experiencing the outcome.[1,2] More insight into the stability of utilities is also important for informing patients about possible future shifts in their evaluations.[3]

Studies that have assessed utilities repeatedly have shown conflicting results. In some groups unstable utilities are found: cancer patients during radiotherapy,[3] breast cancer patients treated with adjuvant chemotherapy,[4] pregnant women,[5] and patients treated for rheumatoid arthritis.[6] In other groups, utilities remained fairly stable: patients receiving hemodialysis,[7] patients with laryngeal cancer undergoing radiotherapy,[8,9] patients undergoing chemotherapy for cancer,[10] and survivors of myocardial infarction.[11]

Several studies of the latter group have concluded that patients' utilities remained stable despite changes in the patients' own clinical states.[8–11] One explanation is a lack of sensitivity of the assessment techniques.[9] Further, a change in preferences may not occur until a later phase in the treatment,[10] or may occur only when the expected side effects of the considered treatment are severe enough to alter preferences.[10] Llewellyn-Thomas et al.[9] observed that patients in the severe-outcome group assigned higher evaluative scores to the severe-outcome sce-

Received September 18, 1998, from the Department of Medical Decision Making (SJTJ, AMS, PPW, JK) and Department of Clinical Oncology (SJTJ, MAN, EMN), Leiden University Medical Center, Leiden, The Netherlands. Revision accepted for publication September 20, 1999. Supported by a grant from the Dutch Cancer Society (Grant 95-734).

Address correspondence and reprint requests to Ms. Jansen: Department of Clinical Oncology, K1P, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands.

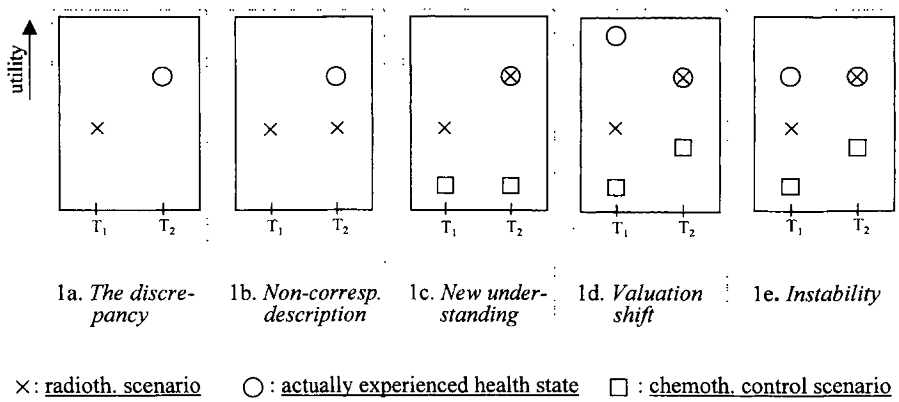

FIGURE 1. The study design for explaining the discrepancy between hypothetical evaluations and actual experience, using assessments before (T_1) and during (T_2) radiotherapy.

nario when in that health state than before. The patients in the mild- and moderate-outcome groups did not change their evaluations. Finally, another explanation for stable utilities may be that preferences that are insensitive to changes in clinical status might be "trait" measures, contrary to "state" measures that are unstable.[8]

Unstable utilities may be explained by new understanding due to experience with the health state. Some studies have shown that patients in a specific health state assigned higher utilities to that health-state scenario than did other groups of respondents not in that health state.[1,12,13] Boyd et al.[1] suggest that the differences they found between patient groups may be explained by the direct knowledge of what life is like with the condition under concern. They suggest that patients may become accustomed to a treatment that they before regarded as highly undesirable. Another explanation for unstable utilities may be that predicted utilities for medical outcomes are inaccurate because people experience difficulties in envisioning unfamiliar situations and because they do not know how they will adapt to the situation.[14–16] Yet another cause for instability may be nonconstant discounting.[5] Values obtained immediately before some event may not represent the long-term values of patients. Finally, instabilities may be found because utilities that are measured after the event or the treatment (retrospective utilities) may be distorted. Patients tend to remember in particular the most intense and the last experience of the episode (peak-end rule), and pay little attention to the duration of the situation.[14–16]

To explain the possible discrepancies between hypothetical and actual evaluations, research into preferences elicited before, during, and after some other event that can cause changes in preferences is warranted. The purpose of our study is to determine whether breast cancer patients' hypothetical evaluations of radiotherapy (predicted utilities) change when they later experience that treatment (actual utilities), and if so, how this discrepancy can be explained. Our study design resembles that of Llewellyn-Thomas et al.,[9] in which three scenarios of varying severity of side effects of radiotherapy were evaluated at the beginning and at the end of treatment. At the end of the therapy the patients were asked to indicate the scenario that best described their own health. Our study differed from that and other studies in that we elicited the utility of radiotherapy by using both a hypothetical radiotherapy scenario and the patients' own experienced health states, at three points in time: before radiotherapy (T_1), during radiotherapy (T_2), and after radiotherapy (T_3). Furthermore, a control chemotherapy scenario was evaluated. We investigated whether there was a discrepancy between a hypothetical scenario, evaluated before the radiotherapy, and the actually experienced health state, evaluated during radiotherapy (figure 1a). When a discrepancy is found, our design enables us to investigate whether it can be explained by (one of) the following four explanations, illustrated by figures 1b to 1e.

Non-corresponding description. A difference between the hypothetical radiotherapy scenario and the actual experience with the treatment. The patient might find that the hypothetical situation does not exactly describe what she experiences during treatment. Thus, the scenario description is non-corresponding. This would be implied by a difference between the hypothetical situation and the actual situation evaluated during radiotherapy, while the evaluation of the radiotherapy scenario between T_1 and T_2 would remain stable (figure 1b).

New understanding. A change in information about the situation, caused by the experience with the treatment. During the radiotherapy, the patient has learned better what radiotherapy entails. Even though the description provided in the scenario is accurate, the way in which the patient values it is different, for instance, because she has learnt to adapt.[1] Thus, the scenario description is adequate, but the valuation changes over time. This explanation would be implied by a higher evaluation of the hypothetical radiotherapy scenario at T_2 than at T_1, while the evaluation of the chemotherapy control scenario would remain stable (figure 1c).

Valuation shift. A change in the valuation of a situation, caused by a change in the status quo, i.e., the actually experienced health state. A poorer actual health state may result in higher valuations of other hypothetical health states ("valuation shift").[17] In our study, this would be implied by higher evaluations of the radiotherapy and the chemotherapy control scenarios at T_2 concomitant with lower evaluations of the actually experienced health state at T_2 than at T_1 (figure 1d).

Instability. A change due to other patient- or disease-related factors, such as random fluctuations, test effects, or time per se. This would be implied by a change in the evaluation of the control chemotherapy scenario (figure 1e).

In summary, we investigated whether discrepancies would occur between predicted and actual utilities, and if so, whether these discrepancies could be explained by: non-corresponding description of the hypothetical health state, new understanding due to experience with the health state, valuation shift due to a new status quo, or instability of preference. The latter three explanations have been discussed in the literature, the first, non-corresponding description, has not. To our knowledge, no design has been used before that could detect it.

Methods

PATIENTS

The subjects were 70 patients diagnosed as having early-stage breast cancer and referred to the department of radiotherapy of the Leiden University Medical Center for a five-to-seven-week course of radiotherapy after lumpectomy or mastectomy. Exclusion criteria were: prior experience with radiotherapy or chemotherapy, a diagnosis of ductal carcinoma in situ, metastatic disease, poor understanding of the Dutch language, and chemotherapy as part of the treatment plan.

The study was introduced to the patients during their first meeting with the radiation oncologist. Just before or shortly after their second meeting with the radiation oncologist but before the start of the treatment, the patients were asked to participate by one of three interviewers.

STUDY PROCEDURES

The scenario descriptions for postoperative radiotherapy and adjuvant chemotherapy were developed on the basis of the literature and of the experiences of medical oncologists, radiation oncologists, and patients. Each scenario contained statements regarding the levels of physical, psychological, and social functioning associated with the health state. Before presenting the description of the actually experienced health state, the patients were asked about their actually experienced health during the preceding week, with explicit reference that health encompasses a physical as well as a psychological and a social dimension. Responses were written down. Each time the scenario was used to elicit utilities, it was emphasized that the patients' actually experienced health state, thus how they had felt during the preceding week, was the subject under concern. The scenario description was used as a reminder only to stimulate the patients to evaluate their health on all three dimensions. Descriptions of the scenarios are given in the appendix.

Utilities for temporary health states were measured by means of a visual analog scale (VAS) anchored by death (0) and perfect health (1), a chained time tradeoff (TTO), and a chained standard gamble (SG), with death (0) and good health (1) as extremes. The choice of the chained methods, introduced by Torrance,[18] was based on the temporary nature of the health states to be evaluated, radiotherapy and chemotherapy. With the chained methods, the temporary health states to be evaluated are weighed not directly against good health and death, but indirectly with the aid of an anchor health-state scenario. The use of the chained procedure requires an anchor health-state scenario that can be evaluated either as a temporary health state, in the first stage of the procedure, or as a chronic health state, in the second stage of the procedure. We opted for a hypothetical "hospitalization, caused by a serious accident," because this situation is one that we felt most patients had probably not experienced but could imagine. We did not use an anchor health-state scenario based on breast cancer, for example metastatic disease, because we considered this too threatening for the patients. To avoid confusion, we decided on a period of six months for all hypothetical scenarios (hospitalization, chemotherapy, radiotherapy) and for the actual health-state scenario. Applied to the radiotherapy scenario, this results in a six-week radiotherapy treatment followed by four and a half months of possible side effects. The quality of life in the periods following the temporary health states had to be the same for all health states. To be realistic, this had to include a prior diagnosis of breast cancer. Therefore, all temporary health states (including good health) comprised surgery for breast cancer in the past, no requirement for further treatment, and a good quality of life.

Using the study design, by means of the chained TTO, the patient first chooses between six months (time t) in the temporary health state to be evaluated (e.g., radiotherapy scenario: Q) versus six months (time x) in the anchor health state (hospitalization scenario: A), both followed by good health for the rest of the patient's life. If the patient prefers the radiotherapy scenario, the time in the anchor health state (time x) is shortened using a ping-pong approach until the patient expresses no preference for either of the two alternatives. At this point the required preference value for the temporary health state is

$$h_Q = 1 - (1 - h_A) \cdot x/t$$

where h is the utility and good health has utility 1.[18]

In the second step of the procedure, the anchor health state is presented as a short-term (three months) chronic health state, followed by death, and is measured by the conventional TTO method. The patient first chooses between three months (time t) in the anchor health state (A) versus six weeks (time y) in perfect health, both followed by sudden and painless death. If the patient prefers perfect health (the anchor health state), the time in perfect health (time y) is shortened (prolonged) until the patient expresses no preference for either of the two alternatives. At this point the required preference value for the anchor health state is[18]

$$h_A = y/t$$

Finally, we get

$$h_Q = 1 - (1 - y/t) \cdot x/t$$

We chose to use the same short-term (three months) time period in the anchor health state for all patients for reasons of comparability of the anchor health-state utilities. This value is the midpoint of the range of the patients' indifference points in the first step of the procedure. The imputation of the anchor health-state utilities in the calculation of the utilities for the health state to be evaluated requires that its utilities, measured in one context with one time duration, can be used in other contexts. Formally, this assumption is based on "utility independence" of the health quality of the anchor health state from life duration (health quality is independent of the duration it is combined with), as well as "separability of preference" over disjoint time intervals (the contribution of the anchor health-state time interval to the overall QALYs is independent of what came before or what comes after). A major improvement of the chained methods is that "utility independence" and "separability" are required only for the anchor health state, which is especially selected to serve this purpose. In conventional QALY measurements, these requirements have to be imposed on all health states. Moreover, we expect that possible violations of these assumptions will occur similarly for all health states and will thus not jeopardize our study results.

By means of the chained SG, the patient chooses between a certainty of six months in the temporary health state to be evaluated (e.g., radiotherapy: Q) and a gamble with a 50% chance of good health for six months and a 50% chance of the anchor health state (hospitalization: A) for six months. All health states are followed by good health for the rest of the patient's life. Depending on the stated preference, the chances of good health (p) and of the anchor health state (1 − p) are varied until the indifference point is reached. At this point the required utility for the temporary health state is

$$h_Q = p + (1 - p) \cdot h_A$$

where h is the utility.[18]

In the second step of the procedure, the anchor health state is presented as a short-term chronic health state and evaluated using the conventional SG method. The length of the anchor health state was set to six months, the duration of all temporary health states, and compared with good health and death. The patient first chooses between a certainty of six months in the anchor health state (A) and a gamble with a 50% chance of good health for six months and a 50% chance of death within one week. All health states are followed by sudden and painless death. Depending on the stated preference, the chances of good health (q) and of the anchor health state (1 − q) are varied until the indifference point is reached. The utility of the anchor health state is[18]

$$h_A = q$$

Finally, we get

$$h_Q = p + (1 - p) \cdot q$$

The imputation of the utility of the anchor health state in the calculation of the utility for the health state to be evaluated is again based on the assumption of "separability of preference," as for the chained TTO measurement.

The patients were interviewed three times: shortly before radiotherapy (T_1), during the final week of radiotherapy (T_2), and two months after radiotherapy (T_3). For answering the research question it is sufficient to compare the evaluations obtained at T_1 and T_2. The measurements at T_3 were added because they offer information about retrospective utilities and give more insight into the (in)stability of utilities.

At T_1, data were collected regarding the date of diagnosis, type of breast surgery, and demographics.

At all measurement points, utilities were elicited for the patients' actually experienced health states, a radiotherapy scenario, a chemotherapy scenario, and the anchor health-state scenario. To control for an order effect, half of the patients rated their actually experienced health states first, followed by the radiotherapy scenario; the other half rated the radiotherapy scenario first. By receiving explicit clues (dimensions) in a hypothetical scenario, patients might be triggered to pay more attention to these dimensions when subsequently valuing their actually experienced health states. On the other hand, by letting the evaluation of the scenario follow the evaluation of the actually experienced health state, patients might be triggered more strongly by the dimensions in the scenario that are similar to their experience. The chemotherapy scenario was always rated last. At T_2 and T_3, evaluations of the four situations were collected in the same order as at T_1. For T_2 the last week of treatment was chosen by the fact that the effects of radiotherapy tend to peak towards the end of therapy. As the patients were asked to rate both their actually experienced health states and the radiotherapy scenario, a distinction could be made between the evaluations of these two situations.

ANALYSES

The means and standard deviations of the utilities were calculated. The chemotherapy scenario was analyzed only for those patients who did not prefer the anchor health-state scenario (the lower anchor) to the chemotherapy scenario. Torrance[18] uses the worst health state as the anchor health state in the chained methods. We preferred, however, to use the same anchor health state for all patients.

To study whether there was an order effect at T_2 (actually experienced health state evaluated first or radiotherapy scenario first), we used independent-samples t-tests. To examine whether there were statistically significant differences between the predicted utilities (the radiotherapy scenario evaluated at T_1) and the actual utilities (the actually experienced health states at T_2), thus the intended discrepancy, pairwise t-tests were used. The possible explanations for the discrepancy were analyzed, using pairwise t-tests, in the following way:

Non-corresponding description. Differences between the utilities for the actually experienced health state and for the radiotherapy scenario, both elicited at T_2 (during radiotherapy), were analyzed. Changes in the utilities of the radiotherapy scenario between T_1 and T_2 were assessed. The explanation is supported if there are differences at T_2 between the utilities for the actually experienced health state and the radiotherapy scenario whereas the utilities for the radiotherapy scenario remain stable over time.

New understanding. Changes in the utilities for the radiotherapy scenario and for the chemotherapy scenario between T_1 and T_2 were tested. The explanation is supported if the utilities for the chemotherapy scenario remain stable whereas the utilities for the radiotherapy scenario show higher values at T_2 than at T_1.

Valuation shift. The stability of the utilities for the

Table 1 • Accrual and Participation of Early-stage Breast Cancer Patients in the Study

Eligible early-stage breast cancer patients entering radiotherapy	109
Declining participation	39 (36%)
First interview	70
Dropout during interview	2 (3%)
Refusal to participate in second interview	10 (14%)
Second interview	58
Dropout during interview	0 —
Refusal to participate in third interview	2 (3%)
Other primary carcinoma	1 (2%)
Third interview	55
Dropout during interview	0 —

Table 2 • Demographics and Clinical Characteristics of the Early-stage Breast Cancer Patients (*n* = 55)

Age	
Median	57 years
Range	33–82 years
Time from diagnosis to first interview	
0 to 1 month	0 —
1 to 2 months	30 (55%)
2 to 3 months	14 (25%)
3 to 4 months	6 (11%)
>4 months	5 (9%)
Prior therapy	
Lumpectomy	49 (89%)
Modified radical mastectomy	6 (11%)
Marital status	
Married/living together	44 (80%)
Widowed	7 (13%)
Single	1 (2%)
Divorced	3 (5%)
Education	
<10 years	27 (49%)
10–15 years	24 (44%)
>15 years	4 (7%)
Occupational status	
Full-time employment	1 (2%)
Part-time employment	21 (38%)
Housewife	33 (60%)

Table 3 • Utilities for Temporary Health State (Scenarios) Elicited from Early-stage Breast Cancer Patients (n = 55) by Means of a Visual Analog Scale, a Chained Time Tradeoff, and a Chained Standard Gamble*

	No.	T_1 Mean (SD)	T_2 Mean (SD)	$\Delta T_1/T_2$ Mean (SD)	p-value $\Delta T_1/T_2$	T_3 Mean (SD)	p-value $\Delta T_1/T_2/T_3$
Visual analog scale							
Actual health state	55	0.81 (0.12)	0.79 (0.14)	0.02 (0.12)	0.17	0.81 (0.13)	0.37
Radiotherapy scenario	55	0.75 (0.13)	0.73† (0.15)	0.02 (0.17)	0.46	0.73 (0.15)	0.67
Chemotherapy scenario	53	0.53 (0.18)	0.49 (0.16)	0.05 (0.18)	0.07	0.51 (0.16)	0.15
Time tradeoff							
Actual health state	54	0.94 (0.12)	0.92 (0.10)	0.02 (0.11)	0.29	0.97 (0.08)	0.02
Radiotherapy scenario	54	0.89 (0.13)	0.91 (0.13)	−0.02 (0.17)	0.31	0.90 (0.17)	0.60
Chemotherapy scenario	25	0.75 (0.24)	0.75 (0.25)	0.00 (0.22)	0.97	0.76 (0.26)	0.93
Anchor health state	52	0.08 (0.14)	0.09 (0.15)	−0.01 (0.18)	0.59	0.08 (0.13)	0.68
Standard gamble							
Actual health state	51	0.94 (0.11)	0.91 (0.19)	0.03 (0.21)	0.30	0.91 (0.20)	0.53
Radiotherapy scenario	51	0.88 (0.19)	0.86† (0.18)	0.02 (0.27)	0.64	0.88 (0.24)	0.86
Chemotherapy scenario	36	0.68 (0.28)	0.75 (0.25)	−0.07 (0.36)	0.24	0.71 (0.28)	0.50
Anchor health state	51	0.27 (0.34)	0.22 (0.24)	0.06 (0.32)	0.19	0.22 (0.29)	0.46

*T_1 = before radiotherapy; T_2 = during radiotherapy; T_3 = after radiotherapy. Stability was studied by pairwise t-tests ($\Delta T_1/T_2$) and analysis of variance for repeated measures ($\Delta T_1/T_2/T_3$).
†Significant difference ($p < 0.01$) between the actually experienced health state and the radiotherapy scenario.

actually experienced health state, the radiotherapy scenario, and the chemotherapy control scenario between T_1 and T_2 was tested. The explanation is supported if the utilities for the radiotherapy and the chemotherapy scenario are higher at T_2 than at T_1, concomitant to lower utilities for the actually experienced health state at T_2 than at T_1.

Instability. The stability of the utilities for the chemotherapy control scenario between T_1 and T_2 was analyzed. This explanation is supported if the utilities for the chemotherapy control scenario change over time.

The mean utilities that were evaluated at T_3 (two months after radiotherapy) are presented. Stability over the three measurement points was analyzed by analysis of variance for repeated measures.

Results

CHARACTERISTICS OF THE PATIENTS

The accrual and the participation of the patients are shown in table 1. Of the 109 eligible patients, 70 patients (64%) participated in the first interview, and 55 of them (79%) completed all three interviews. Characteristics of these 55 patients are given in table 2. The patients were 33 to 82 years old, and the median age was 57. The median time from diagnosis to first interview was between one and two months. Forty-nine patients (89%) had had a lumpectomy and six, a modified radical mastectomy. The majority of patients (80%) were married or were living together with a partner. Half of the patients (49%) had less than ten years of education. The majority (60%) were housewives. Most interviews (90%) were conducted at the patients' home. The 15 patients who did not complete all interviews did not differ significantly in these characteristics. The main reasons for not completing the study were: physical problems (33%), psychological problems (53%), and other problems (13%). The mean utilities at T_1 for these patients were not significantly different from the mean utilities calculated for the other patients.

Forty-eight patients (87%) received a seven-week radiotherapy protocol (33 sessions) and seven patients, a five-week radiotherapy protocol (20–25 sessions). On average, the first interview took place one day before the start of radiotherapy (range: five days before to one day after). Five patients were interviewed between the first and second days of treatment, because there was no possibility to interview them before the first day of treatment. The second interview was, on average, on the 27th day of treatment for the patients with a seven-week course and on the 22nd day of treatment for the patients with a five-week course. On average, the third interview took place on the 59th day after radiotherapy.

THE STABILITY OF UTILITIES

Neither for the actually experienced health state nor for the radiotherapy scenario were order effects found at T_2. This means that the patients (n = 29) who evaluated their actually experienced health states first did not show statistically significant differences in their evaluations from the patients (n = 26) who evaluated the radiotherapy scenario first. The groups were therefore pooled.

In table 3, means and standard deviations of the

Table 4 • Utilities for the Hypothetical Radiotherapy Scenario Evaluated before Radiotherapy and the Actually Experienced Health State during Radiotherapy in Early-stage Breast Cancer Patients (n = 55) by Means of a Visual Analog Scale, a Chained Time Tradeoff, and a Chained Standard Gamble*

	No.	T_1 Mean (SD)	T_2 Mean (SD)	$\Delta T_1/T_2$ Mean (SD)	p-value $\Delta T_1/T_2$
Visual analog scale					
Radiotherapy before versus actual experience	55	0.75 (0.13)	0.79 (0.14)	−0.04 (0.16)	0.07
Time tradeoff					
Radiotherapy before versus actual experience	54	0.89 (0.14)	0.92 (0.10)	−0.04 (0.14)	0.04
Standard gamble					
Radiotherapy before versus actual experience	51	0.88 (0.19)	0.91 (0.18)	−0.03 (0.27)	0.45

*T_1 = before radiotherapy; T_2 = during radiotherapy. Stability was studied by pairwise t-tests ($\Delta T_1/T_2$).

utilities are presented, as well as mean difference scores between T_1 and T_2, their standard deviations, and the p-values for the paired t-test between T_1 and T_2. The last two columns show the mean utilities at T_3 and the p-values for the analysis of variance for repeated measures between T_1, T_2, and T_3.

To study whether there are differences between utilities for a prior hypothetical scenario and an actually experienced health state (the discrepancy), we compared the utilities for the radiotherapy scenario at T_1 (before treatment) with the actually experienced health states at T_2 (during treatment) (table 4). For all three methods, utilities are higher once the state is experienced. By means of pairwise t-tests, a significant difference was found for the TTO [$p = 0.04$, 95% CI = (−0.08, −0.01)]. For the VAS the difference is marginally significant [$p = 0.07$, 95% CI = (−0.08, 0.01)]. Thus, the expected discrepancy was observed. Next, we report the results of testing the various proposed explanations (see figures 2a to 2c and the explanations above).

Non-corresponding description. At T_2 (during radiotherapy) we studied whether the radiotherapy scenario was evaluated differently from the patients'

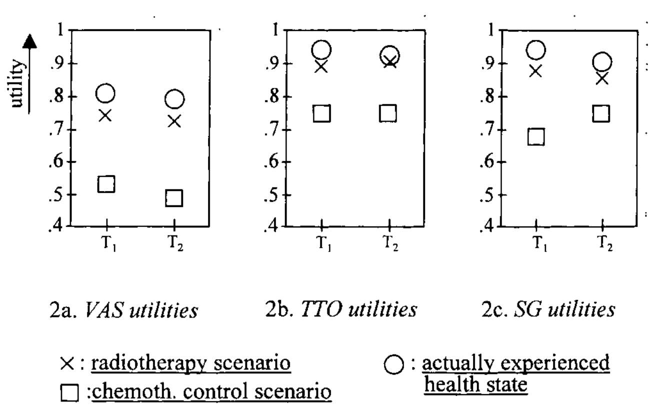

FIGURE 2. Evaluations of (scenarios of) health states before (T_1) and during (T_2) radiotherapy in early-stage breast cancer patients (n = 55).

actually experienced health states. All methods showed higher evaluations for the actually experienced health states. The differences were statistically significant for the VAS [$p \le 0.01$, 95% CI = (0.02, 0.09)] and the SG [$p \le 0.01$, 95% CI = (0.01, 0.08)], but not for the TTO [$p = 0.34$, 95% CI = (−0.02, 0.05)]. The evaluation of the radiotherapy scenario remained stable between T_1 and T_2. These results suggest that our description of radiotherapy did not fully correspond to the experience of the patients during the treatment.

New understanding. The evaluations of the radiotherapy scenario do not show statistically significant changes between T_1 and T_2, and thus the radiotherapy scenario is not evaluated more positively when experienced than when hypothetical. Hence new understanding does not seem to explain the data. The utilities for the chemotherapy control scenario remained stable between T_1 and T_2. Notice that the TTO and the SG chemotherapy utilities were based on 25 and 36 patients, respectively.*

Valuation shift. The actually experienced health states showed lower mean evaluations at T_2 than at T_1 for all methods, although the changes are not statistically significant. These lower evaluations seem natural, because the patients' actually experienced health states did change over time (they were treated with radiotherapy). The evaluations of the radiotherapy scenario and the chemotherapy control scenario do not show statistically significant instabilities. Despite a deterioration in the evaluations of the actually experienced health states during radiotherapy, a "valuation shift"[17] did not seem to occur, because the evaluations of the other health-

*The chemotherapy scenario was analyzed for only those patients who did not prefer the anchor health-state scenario to the chemotherapy scenario. Preceding this study, the chained methods had been pilot tested with ten early-stage breast cancer patients. The chemotherapy scenario was preferred to the anchor health state by almost all patients (90% for TTO and SG), and thus seemed to be an acceptable anchor.

state scenarios remained fairly stable over time. A "valuation shift" would result in higher evaluations of the other health-state scenarios. The phenomenon of valuation shift has been described in the literature for cross-sectional data, i.e., patients in poor health give higher valuations of hypothetical states than do patients in good health. But this may not necessarily imply that a small change in health state over a short period of time will result in such a shift. Indeed, Dolan[17] has suggested that time is needed for the patient to adapt. Besides, it may apply to patients with a severely impaired quality of life only.

Instability. The evaluations of the chemotherapy control scenario do not show statistically significant instabilities with all three methods. These findings suggest that other patient- or disease-related factors did not play a major part in the elicitation of the utilities. It is not likely that the stability of preferences for the treatment scenarios as seen in our study is caused by the patients' recollections of their previous answers, because of the relatively long periods of time between the three measurement points, namely six weeks between T_1 and T_2 and two months between T_2 and T_3.

In the last column of table 3, the utilities obtained two months after radiotherapy are shown. The actually experienced health states show lower mean evaluations at T_2 than at T_1 and higher mean evaluations at T_3 than at T_2. This trend is statistically significant for the TTO ($p = 0.02$), which is mainly due to an increase in the utility for the actually experienced health state between T_2 and T_3 [$p \leq 0.01$, 95% CI = $(-0.08, -0.01)$]. The evaluations of the radiotherapy scenario, the chemotherapy scenario, and the anchor health-state scenario do not show statistically significant instabilities.

Discussion

It has often been suggested that health states may be evaluated more positively when they are experienced than when they are imagined.[1,2,12,13,19] Our results support these suggestions. For all methods, the actual experience was evaluated higher than the predicted utilities. The TTO showed a statistically significant difference. It is, however, not clear from the literature what causes this discrepancy. In our investigation, the discrepancy did not seem to be caused by a global change in preferences over time (new understanding), given the stability of the radiotherapy scenario utilities, or by a change in the actually experienced health state of patients, given the absence of a "valuation shift" effect. Other patient- or disease-related factors (instability) also do not seem to have played a major part, because of the stability of the chemotherapy scenario utilities. The most likely explanation for the observed difference seems to be that the description of the radiotherapy treatment did not fully correspond to the experience of the patients during the treatment (non-corresponding description). Similar explanations may have played a role in the findings of other studies in the literature. The studies that were mentioned in the introduction[8–11] found stable utilities despite changes in health status, when using scenarios. It is possible that the patients in those studies who evaluated scenarios could not recognize themselves fully in the situations described, as is shown in our study, and therefore did not change their preferences even when their own clinical health status had changed.

This implies that it is important to develop hypothetical health-status descriptions that correspond exactly to the perception of patients in the health state under concern, for example, by developing and using guidelines such as those proposed by Llewellyn-Thomas.[20] There is, however, much variation in utilities between individuals, as is shown in this study as well as in other studies (e.g., that of Johnston et al.[21]), and the question remains whether a hypothetical scenario can be developed in such a way that it represents the perceptions of all or most patients in the health state under concern.

When the utilities obtained after radiotherapy were also included in the analysis, no statistically significant change in utilities for the scenarios was found. This supports the observation that the utilities for the radiotherapy scenario and the chemotherapy control scenario remained stable over time. It is remarkable that in the studies in which instabilities of utilities were found,[3–6] utilities were elicited using the actually experienced health states of patients in one way or another, while most studies showing stable utilities used health-status scenarios in some form.[8–10] Our results agree with this observation, in that stable utilities were found using scenarios, and that unstable utilities were found when using the actually experienced health states of patients at T_2.

The feasibility and the consistency of the chained methods have been demonstrated,[22] but the validity of the chained methods has not been examined previously. There are several underlying assumptions of the chained procedure that have not been tested extensively.[21] Examples of such assumptions are that the utility of the anchor health state should not be systematically affected by its duration (utility independence) or by the health state following after (separability of preference). An advantage of the chained method, however, is that these restrictions need to be imposed on the anchor health state only (because the other health states are always associated with the same durations and within the same sequence of health states), and not on all health states as in the traditional QALY model. Jansen et al.[22] and Johnston et al.[21] note that it is important to choose the anchor health state appropriately. These issues point out that more research on the validity of the chained methods is desirable. For now, it seems unlikely that possible violations of the assumptions underlying the method could explain our findings (e.g., by operating in one scenario only and not in another).

Conclusion

We found that utilities for hypothetical health-state scenarios remained stable over time. However, utilities obtained through hypothetical health-state scenarios may not be valid predictors of the value judgments of real-life health states when actually experienced. The finding that a hypothetical scenario was evaluated more negatively than the actual experience with that health state seems to be caused by differences between those situations rather than by a change in evaluation of a same health state (scenario) over time. We hope that future studies of stability of utility will also seek to disentangle the various factors underlying the differences between predicted and experienced utilities, and that these factors will be investigated in other domains.

The authors thank Mrs. I. Overpelt and Mrs. C. Krommenhoek-van Es for their assistance in collecting and processing the data, the radiation oncologists for their support in approaching their patients, and the patients for their contribution to this study.

References

1. Boyd NF, Sutherland HJ, Heasman KZ, Tritchler DL, Cummings BJ. Whose utilities for decision analysis? Med Decis Making. 1990;10:58–67.
2. Tsevat J, Dawson NV, Matchar DB. Assessing quality of life and preferences in the seriously ill using utility theory. J Clin Epidemiol. 1990;43:73S–77S.
3. Llewellyn-Thomas HA, Thiel EC, McGreal MJ. Cancer patients' evaluations of their current health states: the influences of expectations, comparisons, actual health status, and mood. Med Decis Making. 1992;12:115–22.
4. Kestle J, Levine M, Depauw S, Kelleher L. A time tradeoff method for measuring quality of life in stage 2 breast cancer patients. Clin Res. 1989;37:316A.
5. Christensen-Szalanski JJJ. Discount functions and the measurement of patients' values. Women's decisions during childbirth. Med Decis Making. 1984;4:47–58.
6. Bombardier C, Ware J, Russel 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.
7. Canadian Erythropoietin Study Group. Association between recombinant human erythropoietin and quality of life and exercise capacity of patients receiving haemodialysis. BMJ. 1990;300:573–8.
8. Llewellyn-Thomas HA, Sutherland HJ, Ciampi A, Etezadi-Amoli J, Boyd NF, Till JE. The assessment of values in laryngeal cancer: reliability of measurement methods. J Chronic Dis. 1984;37:283–91.
9. Llewellyn-Thomas HA, Sutherland HJ, Thiel EC. Do patients' evaluations of a future health state change when they actually enter that state? Med Care. 1993;31:1002–12.
10. O'Connor AMC, Boyd NF, Warde P, Stolbach L, Till JE. Eliciting preferences for alternative drug therapies in oncology: influence of treatment outcome description, elicitation technique and treatment experience on preferences. J Chronic Dis. 1987;40:811–8.
11. Tsevat J, Goldman L, Soukup JR, et al. Stability of time trade-off utilities in survivors of myocardial infarction. Med Decis Making. 1993;13:161–5.
12. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. J Chronic Dis. 1978;31:697–704.
13. Ashby J, O'Hanlon M, Buxton MJ. The time trade-off technique: how do the valuations of breast cancer patients compare to those of other groups? Quality Life Res. 1994;3:257–65.
14. Kahneman D, Fredrickson BL, Schreiber CA, Redelmeier DA. When more pain is preferred to less: adding a better end. Psychol Sci. 1993;4:401–5.
15. Redelmeier D, Kahneman D. Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. Pain. 1996;66:3–8.
16. Redelmeier DA, Rozin P, Kahneman D. Understanding patients' decisions: cognitive and emotional perspectives. JAMA. 1993;270:72–6.
17. Dolan P. The effect of experience of illness on health state valuations. J Clin Epidemiol. 1996;49:551–64.
18. Torrance GW. Measurement of health state utilities for economic appraisal. A review. J Health Econ. 1986;5:1–30.
19. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC, for the Panel on Cost–Effectivenss in Health and Medicine. The role of cost–effectiveness analysis in health and medicine. JAMA. 1996;276:1172–7.
20. Llewellyn-Thomas HA. Health state descriptions. Purposes, issues, a proposal. Med Care. 1996;34:DS109–DS118.
21. Johnston K, Brown J, Gerard K, O'Hanlon M, Morton A. Valuing temporary and chronic health states associated with breast screening. Soc Sci Med. 1998;47:213–22.
22. Jansen SJT, Stiggelbout AM, Wakker PP, et al. Patients' utilities for cancer treatments: a study of the chained procedure for the standard gamble and time tradeoff. Med Decis Making. 1998;18:391–9.

APPENDIX

Descriptions of the Health-state Scenarios

Health State	Physical	Psychological	Social
Actually experienced	Possibly symptoms or limitations to everyday activities	Emotions, moods	Possibly limitations in work, leisure activities, sport or social contacts
Anchor scenario, hospitalization after a serious accident	Body almost entirely in plaster; barely able to move, get out of bed, wash, dress, or go to the toilet	Anxiety and depression for longer or shorter periods	Being dependent on others for almost everything, e.g., fetching things, for social contact being dependent on those who come to visit
Radiotherapy scenario Daily hospital visit for radiotherapy over a period of six weeks	Skin reactions (warm, red breast and dry skin), general fatigue and listlessness	Feelings of anxiety, worry about one's future health	Limitations to work or other daily activities, restrictions on leisure activities
Chemotherapy scenario During six months one or two hospital visits per three weeks for chemotherapy via an infusion	Nausea, fatigue, hair loss, difficulty in carrying out strenuous activities, frequent need to rest	Dissatisfaction with one's body	Limitations to work or other daily activities, restrictions on social activities