Patient Utilities for Complications Do Not Predict Treatment Choice in Men with Clinically Localized Prostate Cancer

Christopher Saigal, MD, Jeffrey Gornbein, DrPH, and Mark S. Litwin, MD, MPH

• **Objective:** To ascertain whether the utilities assigned by patients to possible complications of prostate cancer treatment predict their actual treatment choice.

• **Design:** Prospective cohort series.

• **Setting:** A university hospital, a Veterans Administration hospital, and a large county hospital.

• **Participants:** 950 men being evaluated for the presence of prostate cancer.

• **Outcome measures:** Time trade-off utilities for current pelvic functions and future pelvic morbidities were elicited using a laptop application. Health-related quality of life was assessed using the Rand 36-Item Health Survey and UCLA Prostate Cancer Index instruments. Treatment choice for localized prostate cancer was recorded.

• **Results:** Utility measurement was feasible and well accepted. A ceiling effect was noted in utilities elicited by the time trade-off method. In multivariate analysis, none of the utilities elicited was a significant predictor of treatment choice.

• **Conclusions:** The time trade-off method of utility assessment seems to function poorly as a measure of patient preferences in the setting of localized prostate cancer. One possible explanation for this finding is that the metric is actually measuring preferences for years of life (ie, the morbidities evaluated are not worth trading away any life span to avoid). This would imply that the time trade-off method is appropriate for morbidities commonly considered to be extremely severe (eg, blindness).

For many men diagnosed with clinically localized prostate cancer, deciding on the optimal treatment is a complex process subject to a variety of influences. The published literature reports apparently similar survival data for many treatments, including radical prostatectomy, brachytherapy, external beam radiotherapy, and, in some series, watchful waiting [1–3]. Because potential side effects of treatment (erectile dysfunction, urinary incontinence, and radiation proctitis) have a significant impact on health-related quality of life (HRQOL), patients often compare the chances of incurring these side effects when making a decision about their care.

One method frequently used to quantify patients’ preferences for health states (such as erectile dysfunction) is utility assessment. A utility is a number (usually between 1.0 and 0) that expresses a patient’s valuation of a health state. The state of perfect health is usually assigned a utility of 1.0 and death is usually assigned a utility of 0. Several methodologies for assessing patient utilities have been described; the standard gamble, time trade-off, and rating scale methods are most commonly employed [4]. One way to demonstrate the validity of a method is to determine whether the utilities elicited by that method have the same relationship to a patient’s treatment choice that the patient’s implicit, or unquantified, preferences have. For example, clinical experience suggests that patients who have a strong aversion to becoming impotent tend to choose treatments that minimize that risk. As such, we would predict that a patient who has a low utility for impotence, as measured by one of the utility-assessment techniques, would also choose a treatment that minimizes the risk of impotence, given that published survival data are similar for the treatment options for localized prostate cancer. We felt that the clinical situation of localized prostate cancer would serve as an excellent setting to determine if time trade-off utilities, held to be part of normative decision-making theory, would also function descriptively in actual decision-making settings.

We designed this study to ascertain the construct validity of the time trade-off method, as evidenced by its ability to elicit utilities that are associated with treatment choice in logical ways. For example, we hypothesized that newly diagnosed patients with low utilities for erectile dysfunction or urinary incontinence would be less likely to choose radical...
PATIENT UTILITIES

prostatectomy, that patients with low utility for radiation proctitis would be less likely to choose external beam radio-therapy, and that patients with low utilities for all treatment-related side effects would be more likely to choose watchful waiting. We also wished to learn whether utility assessment would be well received and practical in busy urology clinics.

Methods

Subjects
Patients scheduled for prostate biopsy between August 1997 and June 2000 were recruited to join the study from 3 clinical settings chosen to ensure a wide racial and socioeconomic mix. These settings included a private academic medical center, a Veterans Administration (VA) medical center, and a county-sponsored hospital.

All English- or Spanish-speaking men scheduled to undergo prostate needle biopsy due to an elevated prostate specific antigen (PSA) level or palpable nodule at any of the institutions were eligible for enrollment. Exclusion criteria included inability to give informed consent or to complete the utility assessment as evidenced by the patient’s irrational ranking of 2 health states, monocular and binocular blindness. Informed consent was obtained according to the guidelines of each institution’s Human Subjects Protection Committee.

Instruments
General HRQOL was measured with the Rand 36-Item Health Survey 1.0 (SF-36) [5]. The SF-36 comprises 8 multi-item scales and has been shown to be both reliable and valid in various populations, including men with prostate cancer [6,7]. A validated cross-cultural Spanish translation of the SF-36 is available [8].

Prostate-targeted HRQOL was measured with the UCLA Prostate Cancer Index (PCI). This self-administered questionnaire quantifies HRQOL specific to prostate cancer in 6 scales and has been shown to be both reliable and valid in populations of older men with and without prostate cancer [9]. A validated cross-cultural Spanish translation of the PCI is available [10].

Sociodemographic and comorbidity data were collected at the time of the survey with a separate instrument that included relevant questions and a 12-item medical history checklist based on an established comorbidity rating scale [11].

Utility assessments were performed with U-Titer II, an established, validated software application that utilizes the time trade-off technique [12]. In brief, the time trade-off technique establishes the amount of life span a patient would “trade away” to avoid having a health problem (eg, impotence). U-Titer II has been used to assess utilities in many clinical conditions, including prostate cancer [13–15]. We assessed utilities for 7 health states: current general health, current urinary, bowel, and sexual function, hypothetical rad-

diation proctitis, stress urinary incontinence, and erectile dysfunction. The scenarios used to describe these health states are presented in the Sidebar.

A cross-cultural Spanish language translation of the U-Titer II program was created for this project. A fluent bilingual research assistant first translated the entire program from English into Spanish. This version was then translated back into English by a fluent bilingual medical interpreter. The 2 English versions were compared, and the 2 translators resolved several minor discrepancies. The revised Spanish translation then underwent pilot testing and cognitive assessment in 8 bilingual volunteer subjects in the county hospital’s urology clinic. A bilingual research assistant went through the Spanish program with the 8 patients by asking each of them to state in English what he was reading in Spanish. No significant conceptual or semantic miscommunications were identified. Several minor grammatical points were clarified in the final Spanish translation.

Data Collection

All interviews were conducted by 1 of 4 project staff members trained in the use of U-Titer II. The interviewer was present during the interview to assist with questions but otherwise remained unobtrusive. Patients were then given the quality-of-life instruments to complete at home and return. At the time of the assessment, serum PSA level and digital rectal examination results were recorded from the chart. Six months after the initial interview, patients were contacted again by a team member. If a review of their chart indicated that they had been diagnosed with prostate cancer at the time of the biopsy, they were asked what, if any, treatment they had undergone.

Statistical Methods

Median utility values for each health state were compared across treatment choice groups. Because the distribution of utility values was not normal, the median was used to represent central tendency, and the Kruskal-Wallis procedure (a nonparametric test) was used to assess for significant differences between the group medians.

The multivariate assessment of the candidate factors for predicting treatment choice was carried out using classification-tree methods [16]. This nonparametric method makes no assumptions about linearity and uses binary recursive partitioning to create a model. Candidate predictors were added to the model if the improvement in the deviance (fit) was significant, using a liberal P < 0.10 criterion. Thus, some overfitting was allowed to reduce the possibility that a significant addition was omitted. The estimated model accuracy, defined as the unweighted average of the sensitivity and specificity of the predicted treatment versus the actual treatment choice, is reported for each model. In multivariate analysis, 19 candidate factors were considered for entry into models for
predicting treatment choice: all 7 utility variables, age, race, marital status, relationship status, education level, number of comorbid conditions, PSA level, digital rectal examination result, Gleason score, and HRQOL scores from the PCI (urinary function, bowel function, and sexual function).

Results
Of 950 men interviewed at baseline, 156 (16%) were excluded because of inability to understand the study task. Of the remaining 794 patients, 214 (27%) had prostate biopsies that revealed adenocarcinoma and chose 1 of 4 treatments: brachytherapy (9%), radical prostatectomy (55%), watchful waiting (20%), or external beam radiotherapy (16%) (Table 1).

Average completion time of the utility assessment was 25 minutes. The median age of the study group was 64 years (range, 40 to 76 years), and 61% of the group were white. The median PSA level was 7.2 ng/mL, and 71% of patients had 2 or fewer comorbid conditions.

Table 2 presents the patient utility data obtained for each health state, stratified by treatment choice. Overall, there were few significant differences among the utilities, and for differences that achieved statistical significance, the actual difference in median utility was of limited clinical relevance. One exception was patients choosing external beam radiotherapy.

About the Time Trade-off Technique of Utility Assessment
The automated interview built with U-Titer II first asks patients to imagine that they are in a state of impaired health (eg, urinary incontinence). The program presents them with a detailed clinical description of the impaired health state. They are then shown a screen and asked to choose between 2 options. One option is to live with the described morbidity for a period of time (corresponding to remaining life expectancy). The second option is to live with their baseline function in the relevant area but live for fewer years. The amount of time remaining in the second option is varied systematically based on the subjects’ choices. The assessment is terminated when the patient is indifferent between the 2 options. Utilities are computed by dividing the number of years the patient has chosen to live in his or her baseline state of functioning by the maximum possible number of years with impaired functioning offered to him or her.

Scenarios Used in Utility Analysis

Incontinence
Imagine that you have DIFFICULTY CONTROLLING YOUR URINE. Imagine that when you cough or sneeze or stand up suddenly, you leak a few drops of urine. Because of this, you need to wear a pad in your underwear. Imagine that you won’t leak a lot, but you will need to wear 1 or 2 pads a day. Think about what you might trade to avoid having this problem.

Impotence
Imagine that you have DIFFICULTY WITH ERECTIONS. Imagine that you have a normal sex drive but cannot get an erection good enough to have sex (either with a partner or by yourself). Think about what you might trade to avoid having this problem.

Proctitis
Imagine that you have DIFFICULTY WITH YOUR BOWELS. Imagine that you sometimes have diarrhea, mucous from the rectum, or a strong urge to move your bowels. Imagine that you have problems like this 2 or 3 times a week. Think about what you might trade to avoid having these problems.

Table 1. Patient Demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study Group (n = 214)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race, %</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>5.1</td>
</tr>
<tr>
<td>African American</td>
<td>28.0</td>
</tr>
<tr>
<td>Latino American</td>
<td>4.2</td>
</tr>
<tr>
<td>White</td>
<td>60.6</td>
</tr>
<tr>
<td>Other</td>
<td>1.9</td>
</tr>
<tr>
<td>Mean ± SD age, yr</td>
<td>62.9 ± 7.2</td>
</tr>
<tr>
<td>Mean ± SD prostate-specific antigen level, ng/mL</td>
<td>10.1 ± 8.6</td>
</tr>
<tr>
<td>Number of comorbid conditions, %*</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Unknown</td>
<td>18</td>
</tr>
<tr>
<td>Treatment choice, n (%)</td>
<td></td>
</tr>
<tr>
<td>Brachytherapy</td>
<td>19 (9)</td>
</tr>
<tr>
<td>Radical prostatectomy</td>
<td>117 (55)</td>
</tr>
<tr>
<td>Watchful waiting</td>
<td>43 (20)</td>
</tr>
<tr>
<td>External beam radiotherapy</td>
<td>35 (16)</td>
</tr>
</tbody>
</table>

*Comorbid conditions: diabetes, myocardial infarction, stroke, amputation, peripheral vascular disease, asthma, emphysema, peptic ulcer disease, renal disease, major depression, seizure disorder, alcoholism, other drug addiction.
whose median utility value for radiation proctitis was significantly higher (more acceptable) than that for patients choosing any of the other 3 treatments.

None of the utilities measured was a significant predictor in any of the multivariate models constructed. For the model constructed to predict choice of watchful waiting (versus all other treatment choices), only 5 factors were found to be significant: age greater than 64 years, increased number of comorbid conditions, marital status (being separated, widowed, or never married), part-time employment status, and urinary function score on the PCI of less than 86. The accuracy of this model was 83%. For the model constructed to predict choice of radical prostatectomy (versus all other treatment choices), marital status (married or divorced), age less than 64 years, PSA level less than 12 ng/mL, and urinary function score on the PCI greater than 86 were the only significant variables. The accuracy of this model was 74%. For the model constructed to predict radical prostatectomy versus the other 2 treatments and excluding watchful waiting (and therefore run on a subset of 171 men), the only significant variables were number of comorbid conditions (greater than 0), age (less than 72 years), and urinary function score on the PCI less than 75. The accuracy of this model was 75%.

**Discussion**

Our study had 2 main objectives. The first was to ascertain the construct validity of the time trade-off method by correlating utilities elicited by this method with ultimate treatment choice. The time trade-off method of utility assessment, as performed in this study, seems to function poorly as a measure of patient preferences for morbidities related to genitourinary and rectal function. For most of the health states measured (with the exception of radiation proctitis), the median utility was near 1.0. This “ceiling effect” indicates that the method fails to discriminate among the gradations in preferences for morbidities, such as urinary incontinence. One possible explanation for this result is that the metric is actually measuring preferences for years of life (ie, the morbidities evaluated are not worth trading away any life span to avoid). This would imply that the time trade-off method is appropriate only for morbidities that are considered extremely severe (eg, blindness). Many health care decisions do not involve risks of extreme disability; the time trade-off metric as administered in this study may be inappropriate in such settings. Although the time trade-off method produced the expected results for radiation proctitis (median utility for radiation proctitis was much higher in men choosing external beam radiotherapy than other treatments), this relationship did not retain significance in multivariate analyses.

Our method of assessing construct validity is limited by the fact that factors besides preference for morbidities will impact a patient’s ultimate treatment choice. Treatment choice can be affected by the relevant experience of family and friends, the professional bias of the treating physician, patient perception of the probability of various morbidities, and factors included in our multivariate analysis such as age, PSA level, and comorbidity. In an observational study of factors that influence women when making the treatment decision to have a hysterectomy, Groff and colleagues found 3 common factors: the perceived outcomes and morbidity of hysterectomy, the perceived views of men/partners, and opinions about health care providers [17]. Because therapeutic options for clinically localized prostate cancer are associated with similar published survival data, it would be
reasonable to extrapolate from subjective expected utility (SEU) theory that preferences for morbidities would be significant drivers of treatment choice. However, if preferences for morbidities are indeed significant predictors of treatment choice, time trade-off utilities, as measured in this study, did not discriminate finely enough among patients to allow us to include them in the predictive model. It is possible that factors unique to our study, such as health state description or patient population, resulted in poor performance of the time trade-off technique.

Our study's second objective was to determine whether it was feasible to perform utility assessment in busy clinical settings using patients who were at risk for the health states being described. We found that the study task was indeed feasible in terms of patient and physician acceptance and time expended per interview. Many patients reported that they enjoyed having explicit attention being paid to their preferences. However, a significant number of patients (16%) were unable to comprehend the time trade-off task well enough to have utilities elicited. The majority of these patients were at the county and VA hospitals. It appears that the cognitive burden associated with the time trade-off method may disproportionately limit its use in some groups of patients.

We could find no studies that prospectively measured utilities using the time trade-off method and correlated them with treatment choice. Stanton and colleagues [18] designed a questionnaire derived from SEU to measure patient preferences in women with breast cancer who were choosing between mastectomy and radical mastectomy. Using this tool, they found that patient preferences for avoiding some of the disfiguring consequences of radical mastectomy predicted choice for breast conserving surgery [18]. That their assessment tool, based on SEU precepts, produced preference values that predicted treatment choice suggests that inherent limitations of the time trade-off method may explain its lack of predictive value in the prostate cancer population. Other utility assessment techniques may perform better in capturing the spread of patient preferences in this population.

Other authors have also found a significant ceiling effect when using the time trade-off method to elicit utilities. Stiggebout examined current health utilities in men with testicular or colorectal cancer [19]. Fully half of each group gave a utility value of 1.0 for their current health. Part of the reason for this skewed distribution may be the result of patient accommodation to chronic health problems (a phenomenon known as response shift) [20]. Nevertheless, the time trade-off technique was insensitive to smaller differences in health state valuation among this population. Stiggebout also varied the number of years in current health available to the patient in the trade-off and found that patients offered longer life spans in their current health were willing to trade more of it away than patients offered shorter life spans. This violation of the “proportionality assumption” (which assumes that utility scores do not depend on the amount of life span offered in the time trade-off assessment) invalidates the calculation of the time trade-off utility score. We offered periods of time to our subjects based on actuarial estimations of their remaining life spans; it is possible that their utility ratings would have varied with offers of shorter times.

In an effort to distinguish the comparative validity of the 3 major utility elicitation techniques, Giesler and colleagues [21] proposed 2 metrics by which to measure the strengths of the different techniques. They proposed a differentiation index, indicating how well the utility technique distinguishes different health outcomes from one another, and an inconsistency index, indicating how often a technique indicates a preference order for health states inconsistent with explicit patient ranking. As in our study, which produced similar utility scores for subjects’ current sexual and urinary function and hypothetical impaired sexual and urinary function, Giesler found that the time trade-off method scored poorly in its ability to differentiate health states along a continuum. In that study, the rating scale method performed best among all those tested.

One of the goals of health services research is to understand the drivers of health resource utilization and to promote rational allocation of health resources. Decision-analysis techniques offer the promise of helping both doctor and patient come to rational decisions using medical evidence that optimize patient utility. Employing valid and appropriate methods of utility assessment is critical to this endeavor. Our study did not confirm the validity or clinical usefulness of the time trade-off technique. Other utility assessment techniques may prove to be more useful. However, as Nease [22] has observed, SEU need not be descriptive of actual patient decisions in order to be normative. Additionally, the process of utility elicitation alone may be useful for patients who are anxious to consider all of the quality-of-life implications of their treatment decisions.

Corresponding author: Mark S. Litwin, MD, MPH, UCLA Department of Urology, Box 951738, Los Angeles, CA 90095-1738, mlitwin@ucla.edu.

Funding/support: Financial support for this study was provided entirely by a grant from the American Cancer Society (Research Project Grant No. 98-100-01-PBP). The funding agreement ensured the authors' independence in designing the study, interpreting the data, and writing and publishing the report.

Financial disclosures: None.

Author contributions: conception and design, CS, MSL; analysis and interpretation of data, CS, MSL; drafting of the article, CS; critical revision of the article for important intellectual content, CS, MSL;
final approval of the article, CS, MSL; provision of study materials or patients, MSL; statistical expertise, JG; obtaining of funding, CS, MSL; collection and assembly of data, CS.

References


