American Journal of Gastroenterology
© 2005 by Am. Coll. of Gastroenterology
Published by Blackwell Publishing

ISSN 0002-9270
doi: 10.1111/j.1572-0241.2005.40588.x

Patient-Derived Health State Utilities for Gastroesophageal Reflux Disease

Lauren B. Gerson, M.D., M.Sc., Nighat Ullah, M.D., Trevor Hastie, Ph.D., George Triadafilopoulos, M.D., and Mary Goldstein, M.D., M.S.

Division of Gastroenterology and Hepatology, Stanford University School of Medicine, Stanford, California; Department of Statistics, Stanford University, Stanford, California; VA Palo Alto Health Care System, Palo Alto, California; Geriatrics Research Education and Clinical Center (GRECC), VA Palo Alto Health Care System, Palo Alto, California; and Center for Primary Care and Outcomes Research, Department of Medicine, Stanford University School of Medicine, Stanford, California

BACKGROUND AND AIMS: Gastroesophageal reflux disease is a chronic disease that adversely affects health-related quality of life. The purpose of this study was to derive health state utilities for patients with chronic heartburn symptoms.

METHODS: We used a custom-designed computer program in order to elicit utilities with the time-tradeoff and standard-gamble techniques. Patients with chronic (more than 6 months) symptoms of gastroesophageal reflux disease entered the study. Two interviews were performed in random sequence either initially on medications for heartburn that adequately controlled symptoms, or off of medications for 1 wk while the patient was symptomatic. We also collected data using visual-analog scales, quality of life in reflux and dyspepsia (QOLRAD), and Gastrointestinal Symptom Rating Scale (GSRS) scores.

RESULTS: We invited 222 patients to participate; 158 (71%) patients (129 men, 29 women) completed the study. Barrett's esophagus was present in 40 (25%), erosive disease in 17 (11%), and 118 (74%) had comorbid conditions. The mean (±SD) utility ratings were 0.94 ± 0.09 on medical therapy and 0.90 ± 0.12 off medications for patients with reflux alone using time tradeoff ($p = 0.004$), and 0.94 ± 8.0 both on and off of antireflux medications with standard-gamble assessment ($p = 0.96$). Mean time-tradeoff scores were also significantly lower off of medications for patients with other comorbid conditions ($p = 0.002$). There was no significant difference between mean utility scores for patients with or without Barrett's esophagus or erosive disease.

CONCLUSION: Gastroesophageal reflux disease adversely affects health-related quality of life. Time-tradeoff utility for patients with reflux disease is substantially higher when patients are on medication than off medications.

(Am J Gastroenterol 2005;100:524–533)

INTRODUCTION

Gastroesophageal reflux is a chronic condition affecting 44% of all Americans at least once per month (1), 14% at least once per week (2), and up to 7% daily (3). Administration of the Medical Outcomes Survey (SF-36) to patients with chronic reflux has shown that patients with untreated heartburn symptoms score lower than U.S. population normal controls on all eight subscales of the Medical Outcomes Study SF-36 (4), and that these scores return to normal with medical therapy that adequately controls reflux symptoms (5).

There are many methods available to measure health-related quality of life including generic instruments, disease-specific questionnaires, and patient-derived utility measurements. Generic and disease-specific instruments are useful for the assessment of change over time and in order to compare patients with different diseases. Examples of generic instruments that have been applied to patients with heartburn include the Medical Outcomes Study SF-36 (6) and the Psychological General Well-Being Index (7). Disease-specific instruments for gastroesophageal reflux disease include the quality of life in reflux and dyspepsia (QOLRAD) (8) and the gastroesophageal reflux disease health-related quality of life (GERD-HRQL) (9). Health-state utilities are preference-weighted measures required for cost-effectiveness analyses (10). Utility measurements are designed to provide a numerical indicator of symptom severity based on a scale of 0 to 1.0, where 0 represents death or a state equivalent to death, and 1.0 represents ideal health, or a state without a particular disease such as heartburn. A paucity of health state utilities data exists for patients

with chronic gastroesophageal reflux disease and Barrett's esophagus.

The purpose of this study was to derive health state utilities for patients with chronic reflux symptoms who required daily medication for adequate symptom control. We hypothesized that the reported utility values would change depending upon medication status at the time of the interview and in the presence of other comorbid conditions, including Barrett's esophagus.

SUBJECTS AND METHODS

The study was conducted at Stanford University Hospital and the VA Palo Alto Health Care System. We recruited patients undergoing evaluation for heartburn symptoms in 2002–2003 from the Gastroenterology Clinics, Endoscopy Suites, and Motility Laboratories at both institutions. In addition, patients with known gastroesophageal reflux disease or Barrett's esophagus who underwent evaluation between 1998–2001 were sent a letter of invitation explaining the purpose of the study. Interested patients were instructed to contact one of the investigators or to return the letter indicating permission to be contacted. Potential subjects were interviewed by one of the investigators prior to upper endoscopy, 24-hr ambulatory pH testing, or esophageal motility assessment to confirm that they had typical symptoms of heartburn and/or regurgitation for at least 6 months requiring daily medical therapy for symptom control. Patients were also required to have an increase in their reflux symptoms upon temporary cessation (1–3 days) of medical therapy. This information was either elicited by patient history, or patients were asked to stop their medications for reflux temporarily and to call the study coordinator if they had a change in symptoms that would qualify them for enrollment into the study. Patients with typical reflux symptoms and normal endoscopic exams were required to undergo 24-hr ambulatory esophageal pH studies in order to confirm the diagnosis of pathologic gastroesophageal reflux disease as part of the medical protocol for their management. Exclusion criteria included the presence of a terminal illness, such as unresectable cancer, inability to interact with the computer program, or inability to provide informed consent.

Two interviews were performed in a sequence either initially on medications for heartburn that adequately controlled symptoms, or off medications for 1 wk while the patient was symptomatic, depending on whether the patient was on or off antireflux medications. Patients who were on their reflux medications during the initial interview were asked to return for the second session after cessation of their medications for 7 days. Alternatively, patients who were interviewed initially off reflux medications (usually in the setting of a 24-hr ambulatory pH monitoring study) were asked to return for a repeat session after resuming medical therapy that adequately controlled their reflux symptoms for at least 14 days. All interviews were conducted by either Dr. Gerson or the study coordinator. The study coordinator was trained using a previously utilized instruction protocol developed for NIH/NIA R01 AG15110 and was supervised during the initial 10 patient interviews. Neither the patient nor the interviewer was blinded to the medication status of the patient at each interview.

During endoscopy, the appearance of the esophagogastric junction (defined as the junction of the proximal gastric folds and the tubular esophagus) was carefully studied in the antegrade view, and after retroflexion of the endoscope in the stomach. The squamo-columnar or neo-squamo-columnar junction was also identified as the point where the squamous mucosa joined the salmon-colored columnar mucosa of the Barrett's mucosa. All subjects were classified as having either a normal-appearing esophagogastric junction or the presence of Barrett's esophagus if endoscopic biopsy from an area of salmon-colored mucosa above the esophagogastric junction revealed the presence of specialized intestinal metaplasia.

The study was conducted under approval from the Human Subject Panel for Stanford University and the VA Palo Alto Health Care System.

UTILITY AND QUESTIONNAIRE ASSESSMENT

A computer program for assessing utilities was used to facilitate standardization of the interview. The program, Utility Assessor, was developed by the authors in accordance with standard utility methods (11–13). Once the age of the patient was entered into the first screen, the program calculated life expectancy based upon U.S. Life Tables (14) for use in the subsequent time-tradeoff utility questions. Any patient with a chronic disorder requiring more than 6 months of medical therapy was considered to have a comorbid condition.

Patients initially provided assessments for their current health state. Using a visual-analog rating system, patients were shown a scale where 0 was the worst possible health state or death, and 100 represented a state of perfect health without heartburn or any other problems. The cursor started in the middle of the bar at 50, and was displaced by the research assistant according to the patient's instructions until the patient arrived at a decision about the current value of his/her health. We used the visual-analog scale as a potential predictor for time-tradeoff and standard-gamble scores.

Patient-derived health-state utilities were elicited by the time-tradeoff and standard-gamble techniques. Patients were instructed to consider only typical symptoms of heartburn and/or acid regurgitation and not other symptoms including chest pain, dysphagia, cough, hoarseness, or any other extraesophageal manifestations. During the time-tradeoff interview, subjects were presented with the average number of remaining life years calculated from the life tables (15) according to the age entered into the program. Subjects were then asked how many years or months they were willing to trade in exchange for complete resolution of heartburn symptoms. Patients were allowed to trade no time, as little as 1 month, or up to the maximum number of remaining life

years. For the standard-gamble assessment, patients were told to imagine that there was a theoretical cure for their heartburn that would carry a small risk of a painless immediate death. Subjects were then asked to engage in a gamble between the chance of a cure and risk of death until they found a risk level where they were ambivalent between the two health states. The computer program used a recursive "ping-pong" algorithm in order to arrive at an unbiased answer (16). During the experiment, the research associate was instructed not to offer comments that would bias the subjects in terms of their answers. However, patients who chose not to trade away any of their life were questioned by the assistant in order to insure comprehension of the question. Subjects who did not want to gamble any of their life were shown a subsequent screen inquiring whether the risk of death was truly 0% or were given the option of entering a numerical value between 0–1%. In order to train patients how to respond to the utility questions, we administered practice questions using the time-tradeoff and standard-gamble techniques with a hypothetical scenario of vision loss (17). The subject was instructed to imagine a state of blindness with 10 remaining years to live and was asked how much of his life he would be willing to trade or gamble in order to restore perfect vision. The calculated utility value was displayed to the patient after each practice question. The time-tradeoff value was equivalent to 1 minus the number of years traded divided by the number of life years remaining. The standard-gamble response was equivalent to 100% minus the percentage risk of death that the subject would be willing to accept, converted to a scale from 0 to 1. Patients who did not understand the practice questions received further training by the research assistant who did not advance to the actual utility scenarios until comprehension was obtained. All patients in our current study were able to understand the task and proceed to the actual rating.

We subtracted any decrement in quality of life due to other comorbid conditions in order to derive the decrement in health-related quality of life attributable only to heartburn symptoms (18, 19). Barrett's esophagus was considered as part of the spectrum of GERD and not alone as a comorbid condition. Patients were initially asked how much of their life they would trade in exchange for all of their health problems, including reflux symptoms, and then asked how much they would exchange for all of their conditions other than heartburn. A patient who had no comorbid conditions would not be expected to trade any life years during the second time-tradeoff session. For example, a 55-yr male with heartburn, asthma, and hypertension would be asked how much of 19 yr he would trade to be restored to perfect health. If the patient elected to trade 10 yr of his life in order to exchange his reflux and other problems for perfect health, and then 5 yr to trade the asthma and hypertension, his time-tradeoff utility would be $[1 - (10 - 5)/19]$ or 0.74. During the standard-gamble assessment if the patient chose a 10% risk of death in exchange for his heartburn and other conditions, and a 5% risk of death for the asthma and hypertension without the reflux symptoms, then the utility by standard gamble would equal $1.0 - (0.1 - 0.05)$ or 0.95. Patients who would not trade any of their life or take any risk of death would have a utility of 1.0, or a utility equivalent to perfect health. A subject with a serious comorbid condition, such as a cancer, might attribute very little weight to reflux symptoms, and trade the same amount of years for both questions, again recording a 1.0 score for the time-tradeoff scenario.

All subjects were also asked to complete the Gastrointestinal Symptom Rating Scale (GSRS) and the QOLRAD survey both on and 7 days off of their heartburn medications. Patients received $25 for completion of the first interview and questionnaire completion and $50 for the second interview and questionnaire assessment.

STATISTICAL ANALYSIS

The clinical characteristics of gastroesophageal reflux disease patients with and without Barrett's esophagus were compared. Chi-squared tests were used to analyze categorical variables using SAS (version 6.12, SAS Institute, Cary, NC). Paired Student's *t*-tests were used in Microsoft Excel 2000 (Microsoft Corporation, Redmond, WA) to analyze continuous variables. The level of significance was set at $p < 0.05$. All tests were two-tailed.

We used linear regression to analyze the relationship between the time-tradeoff and standard-gamble values and potential predictors. Mixed-effects linear regression was performed for time-tradeoff and standard-gamble scores. Initial histograms indicated that these response data were peaked at 1.0 (perfect health), and skewed to the left of 1.0. Since the utility scores are numbers that are skewed and bounded between 0 and 1, we applied the "logit" transformation. The started logit transformation of the time-tradeoff and standard-gamble scores made these distributions appear more Gaussian (although the peak on the right remains). We used "started" logits (20) in order to deal with responses that were exactly 1.0 using the formula $\log(p + c)/(1 - p + c)$ where p is proportion, and $c > 0$ is a small constant (1/200).

Since each subject appeared twice in the dataset, both on and off medications, observations were correlated in pairs (21). To account for this correlation, we introduced a random subject effect, which allowed for a different level for each subject. We assumed that the random subject effect is Gaussian with mean 0 and unknown variance. We used a mixed-effects model to estimate all of the fixed effects, as well as the variance of the random subject effect by Gaussian maximum likelihood. Potential predictor variables included medication status, patient age, gender, site of care at Veterans Affairs Medical Center, presence of Barrett's esophagus and/or comorbid conditions, duration of heartburn symptoms, and years on medications. We created scatterplots in order to determine a potential correlation between utility values and questionnaire scores.

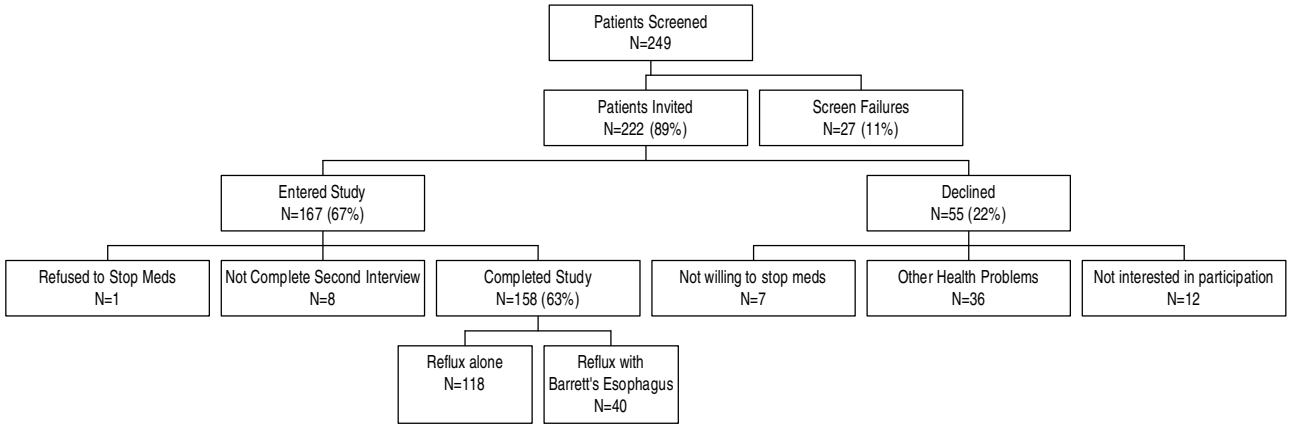

Figure 1. Subject enrollment flow diagram. Patients were invited either before a scheduled appointment in the endoscopy suite or motility laboratory, or from a prior database of patients with gastroesophageal reflux disease and/or Barrett's esophagus. All patients were sent a letter of invitation to participate, and were not required to state the reason for nonparticipation. One patient refused to stop antireflux medication after the initial interview, and eight patients did not complete both interviews.

SAMPLE SIZE CALCULATIONS

Based upon preliminary pilot data in 10 patients with reflux, the mean utility was 0.95 with a SD of 0.10. We excluded the pilot data from the final analysis. In order to detect a potential difference in utility values between patients with and without comorbid conditions and/or Barrett's esophagus of 10% or more (a change in utility of approximately 0.1) with two-tailed $\propto$ of 0.05 and a power of 90%, we calculated that 30 patients would be needed for each group. Assuming that 25% of patients would not return for the second interview, we would require at least 38 patients each with reflux alone, reflux with comorbid conditions, and Barrett's esophagus.

RESULTS

Two hundred and forty-nine patients were screened and 222 were invited to participate in the study (Fig. 1). Twenty-seven patients were screen failures for the following reasons: 10 had a prior Nissen fundoplication and were no longer symptomatic from reflux; 2 did not have gastroesophageal reflux by pH monitoring study; 1 patient was found to have achalasia; 6 patients no longer required heartburn medications for their symptoms; 1 was unable to complete the interviews due to language barrier; 3 patients lived out of state and could not complete the interviews in person; and 4 did not have daily heartburn. Fifty-five (22%) patients declined participation. Most patients declined participation due to the presence of other comorbid conditions requiring active medical attention (n = 36). The other patients stated that they were not interested (n = 12), or were not willing to stop heartburn medications for the study (n = 7).

Of the 158 patients completing both interviews, 118 (75%) had reflux alone, 19 (12%) had erosive esophagitis, and 40 (25%) had Barrett's esophagus in addition to chronic gastroesophageal reflux disease. The mean (±SD) DeMeester score on 24-hr ambulatory esophageal pH monitoring was 40.4 ± 26 (range 15–121.5) for the cohort with nonerosive disease on upper endoscopy.

The baseline demographics for the participating patients are shown in Table 1. The mean (±SD) age for the group was 55 ± 0.8 yr (range 21–87) and was significantly older for the patients with Barrett's esophagus ($p = 0.009$). As expected, the patients with Barrett's esophagus were more likely to be males ($p = 0.04$) with a longer duration of reflux symptoms compared to the subjects with reflux only ($p = 0.03$). While 88 of 159 (55%) of the overall study population were patients seen at the Veterans Affairs facilities, they represented 77% of the patients with Barrett's esophagus. Self-reported comorbid conditions were present in 119 (75%) of the patients and

Table 1. Baseline Characteristics of the Study Population*

Variable	Patients with Reflux Alone (N = 118)	Patients with Barrett's Esophagus (N = 40)	p-Value†	Patients with Comorbid Conditions‡ (N = 119)	Patients without Comorbid Conditions‡ (N = 39)	p-Value†
Age	53 ± 14 (21–87)	59 ± 12 (40–79)	0.009	57 ± 13 (28–87)	45 ± 14 (21–75)	0.000009
Male gender	91/118 (77%)	37/40 (93%)	0.04	98/119 (82%)	30/39 (77%)	0.65
Years with reflux	11.5 ± 12 (0.5–60)	16 ± 12 (1–50)	0.03	15 ± 14 (0.5–60)	8 ± 7 (0.5–30)	0.00003

*Mean ± SD.
†Student's t-test. Chi-square test used for gender analysis.
‡The same patient cohort is analyzed twice in the Table. The two left columns demonstrate the findings according to underlying diagnosis (patients with and without Barrett's esophagus), while the right two columns show the analysis when the same patients are stratified according to the presence or absence of other comorbid conditions.

included hypertension in 54 (34%) patients, cardiac disease in 29 (18%), peptic ulcer disease in 30 (19%), pulmonary disease in 23 (14%), diabetes mellitus in 18 (11%), asthma in 25 (22%), sinus condition in 31 (19%), chronic cough in 17 (11%), arthritis in 56 (35%), depression in 42 (26%), anxiety disorder in 28 (18%), prior cancer in 24 (15%), and chronic pain in 51 (32%). Other less common disorders included epilepsy, anemia, liver disorders, immunodeficiency, stroke, glaucoma, renal disorders, sexually transmitted disease, and other psychological disorders.

Most (131/155 or 85%) of the patients required proton pump inhibitor therapy for symptom control, with 52% of the patients requiring once daily dosing and 48% twice daily therapy. All of these patients reported increase in symptoms on reduced dosage of medication. Six percent of the patients were taking H_2-receptor antagonists with adequate symptom relief, while 9% required H_2-receptor antagonists in addition to twice daily proton pump inhibitors.

The mean (±SD) values for time tradeoff and standard gamble are shown while patients were on and off of antireflux therapy are shown in Table 2. Subjects were more willing to trade or gamble in exchange for reflux plus their other conditions compared to their other health problems without reflux symptoms. The mean risk of death in the standard gamble was 0.12 ± 0.0007 (range 0–0.82) when patients were asked to gamble reflux plus their other problems, and was 0.07 ± 0.006 (range 0–0.87) for the other health problems without reflux. Forty-eight (31%) patients would not trade any time during the initial time-tradeoff scenario, compared to 14 (9%) patients on the first standard-gamble question.

The mean ± SD utility value for the patients' own current health using the time-tradeoff technique was 0.94 ± 0.09 for patients with reflux alone while on antireflux therapy, compared to 0.909 ± 0.12 while off of antireflux therapy ($p = 0.004$, Table 2). There was no significant difference between the standard-gamble scores, however, when patients were off of antireflux therapy. Medication status did not affect utility scores in patients with Barrett's esophagus. The time-tradeoff score was significantly lower off of medications in patients with other comorbid conditions, but no significant differences were seen in patients with erosive esophageal disease.

Most of the patients (142 or 90%) were interviewed initially off of antisecretory medications because they were undergoing evaluation for gastroesophageal reflux disease at the time of the initial interview. For the 16 patients interviewed initially on antireflux medications, the mean ± SD time-tradeoff values was 0.94 ± 0.05 and was 0.92 ± 0.08 off of medications (mean difference 0.028, $p = 0.24$). The mean time tradeoff for the 142 patients who were initially off of medications was 0.90 ± 0.12 compared to 0.93 ± 0.90 when the cohort was on medical therapy ($p = 0.63$). In order to compare whether there was a significant difference between the two groups, we compared the mean differences ± SDs and found no significant difference in time-tradeoff values depending upon initial medication status (mean difference of 0.28 ± 0.08 for time tradeoff in the on-to-off group and 0.037 ± 0.9 for time tradeoff in the off-to-on group, $p = 0.92$).

In order to assess whether there would be a difference between time-tradeoff and standard-gamble assessments, we

Table 2. Health-State Utility Assessments According to Medication Status*

Study Group	Utility Scores On Medication	Off Medication	*p*-Value
Patients with reflux alone, time tradeoff (N = 118)	0.94 ± 0.09 (0.5–1.0)	0.90 ± 0.12 (0.3–1.0)	0.004
Patients with reflux alone, standard gamble (N = 118)	0.94 ± 7.7 (0.6–1.0)	0.94 ± 8.9 (0.25–1.0)	0.96
Patients with Barrett's esophagus, time tradeoff (N = 40)	0.92 ± 0.08 (0.73–1.0)	0.90 ± 0.12 (0.45–1.0)	0.39
Patients with Barrett's esophagus, standard gamble (N = 40)	0.95 ± 0.05 (0.8–1.0)	0.93 ± 0.07 (0.6–1.0)	0.16
Patients with comorbidity, time tradeoff (N = 119)	0.94 ± 0.08 (0.6–1.0)	0.90 ± 0.11 (0.33–1.0)	0.002
Patients with comorbidity, standard gamble (N = 119)	0.89 ± 0.13 (0.52–1.0)	0.93 ± 0.13 (0.25–1.0)	0.28
Patients with erosive esophagitis, time tradeoff (N = 19)	0.93 ± 0.08 (0.75–1.0)	0.87 ± 0.18 (0.35–1.0)	0.19
Patients with erosive esophagitis, standard gamble (N = 139)	0.96 ± 0.08 (0.65–1.0)	0.95 ± 0.05 (0.8–1.0)	0.65
On-to-off medication status, time tradeoff (N = 16)	0.95 ± 0.05 (0.86–1.0)	0.92 ± 0.08 (0.733–1.0)	0.24
Off-to-on medication status, time tradeoff (N = 142)	0.90 ± 0.12 (0.33–1.0)	0.94 ± 0.90 (0.5–1.0)	0.63

*Values shown are mean ± SD. Numbers in parentheses are the ranges for each utility value. Table shows mean utility values on and off medications for the 158 subjects stratified according to time-tradeoff or standard-gamble assessment, presence of comorbid conditions, Barrett's esophagus, or erosive esophageal disease, or order that the interview occurred.

Table 3. Results Comparing Time-Tradeoff to Standard-Gamble Scores*

Study Group	Nonutility Value Visual-Analog Scale	Utility Scores Time-Tradeoff Score	Standard-Gamble Score	p-Value†
Patients with reflux alone on therapy (N = 118)	73 ± 18 (15–100)	0.94 ± 0.09 (0.5–1.0)	0.94 ± 0.08 (0.6–1.0)	0.9
Patients with reflux alone off therapy (N = 118)	63 ± 22 (20–98)	0.90 ± 0.12 (0.3–1.0)	0.94 ± 0.09 (0.25–1.0)	0.003
Patients with Barrett's esophagus on therapy (N = 40)	69 ± 19 (10–100)	0.92 ± 0.09 (0.62–1.0)	0.95 ± 0.05 (0.8–1.0)	0.05
Patients with Barrett's esophagus off therapy (N = 40)	55 ± 21 (15–99)	0.90 ± 0.12 (0.45–1.0)	0.93 ± 0.07 (0.6–1.0)	0.24
Patients with comorbidity‡ (N = 119)	59 ± 23 (15–100)	0.90 ± 0.11 (0.33–1.0)	0.94 ± 0.07 (0.6–1.0)	0.006
Patients without comorbidity‡ (N = 39)	67 ± 21 (20–98)	0.89 ± 0.13 (0.52–1.0)	0.93 ± 0.13 (0.25–1.0)	0.28
Patients with erosive esophagitis‡ (N = 17)	60 ± 6 (20–90)	0.92 ± 0.03 (0.6–1.0)	0.94 ± 0.01 (0.8–1.0)	0.46

*Values shown are mean ± SD. Numbers in parentheses are the ranges for each value.
†Student's t-test comparing mean time-tradeoff and standard-gamble values.
‡Mean values off of medications for reflux symptoms.

compared the mean utility values for patients with reflux alone and Barrett's esophagus both on and off of antireflux therapy as well as for patients with other comorbid conditions and erosive disease (Table 3). Off antireflux therapy the time-tradeoff scores were significantly lower compared to standard-gamble scores for patients with gastroesophageal reflux disease without Barrett's esophagus, and in the patients with other comorbid conditions.

The scatterplot matrix in Figure 2 shows all pairwise plots of the logit-transformed values for VAS and the utility values, as well as scores on the symptom surveys. Scatterplots of the data (Fig. 2) demonstrated a correlation between the scores obtained using the visual-analog scale and the mean utility values, but no correlation between the utility scores and the scores from the disease-specific surveys. The mean scores for the reflux and abdominal pain subscales of the GSRS and the quality in life in reflux and dyspepsia food and drink subscales changed significantly when patients were on and off medications for reflux (mean reflux score 2.8 ± 1.5 on medications and 3.9 ± 1.7 off medications, $p < 0.0001$; mean abdominal pain score 2.5 ± 1.3 on medications and 3.2 ± 1.6 off medications, $p = 0.0001$; and mean food and drink score 5.1 ± 1.5 on medications and 4.2 ± 1.6 off medications, $p = 0.000005$). However, none of the survey scale scores changed when comparing patients with and without Barrett's esophagus and with and without comorbid conditions.

Separate regression analyses were performed for the time-tradeoff scores and the standard-gamble scores using the overall mean time-tradeoff score or standard-gamble score as the dependent variable, and the presence or absence of medical therapy, Barrett's esophagus, and/or comorbid conditions as the independent variables. The visual-analog scores were highly correlated with the utility values and were dropped from both models.

For the time-tradeoff score, the full-fitted model (mean values ± SEM) was:

$$\text{Log}(p/(1-p)) = (2.4 \pm 0.43,\ p < 0.001) + (0.5 \pm 0.11 \times \text{medication status},\ p < 0.001) + (0.02 \pm 0.009 \times \text{years of symptoms},\ p = 0.04), -(0.3 \pm 0.27 \times \text{veteran status}) - (0.4 \pm 0.23 \times \text{presence of Barrett's esophagus},\ p = 0.12) - (0.05 \pm 0.29 \times \text{gender},\ p = 0.86) + (0.13 + 0.01 \times \text{age},\ p = 0.14) - (0.16 + 0.28 \times \text{presence of cormorbid conditions},\ p = 0.57) - (0.02 \pm 0.03 \times \text{years on medications},\ p = 0.54).$$

This model suggested that a higher score on the time-tradeoff exercise was associated with being on medication at the time of the interview and a greater duration of heartburn symptoms. Since the responses were transformed using the started logit transformation, the coefficients of the model estimate log-odds-ratios. When a mixed-effects model was performed for the standard-gamble score, none of the predictors were found to be significant.

DISCUSSION

Gastroesophageal reflux disease is a chronic condition affecting millions of Americans on a regular basis. In this prospective study, we have defined the health-state utilities associated with reflux symptoms and have demonstrated that the impact on health utility is significant. Off reflux medications, the time-tradeoff values approached the values obtained for patients with other serious conditions, such as angina pectoris (22), New York Heart Association Class I or II congestive heart failure (23), coronary heart disease (24), chronic obstructive pulmonary disease (25), and asthma (26).

It is important to note that all patients in the study had chronic daily heartburn symptoms requiring medical therapy for symptom relief. We felt that 1 wk off of medications was sufficient for patients to return to "off-medication" symptom levels since patients were required to have increased reflux symptoms upon temporary cessation of antireflux therapy for 1–3 days prior to entry into the study. In addition, most patients would not accept being off of antireflux therapy for more than 1 wk.

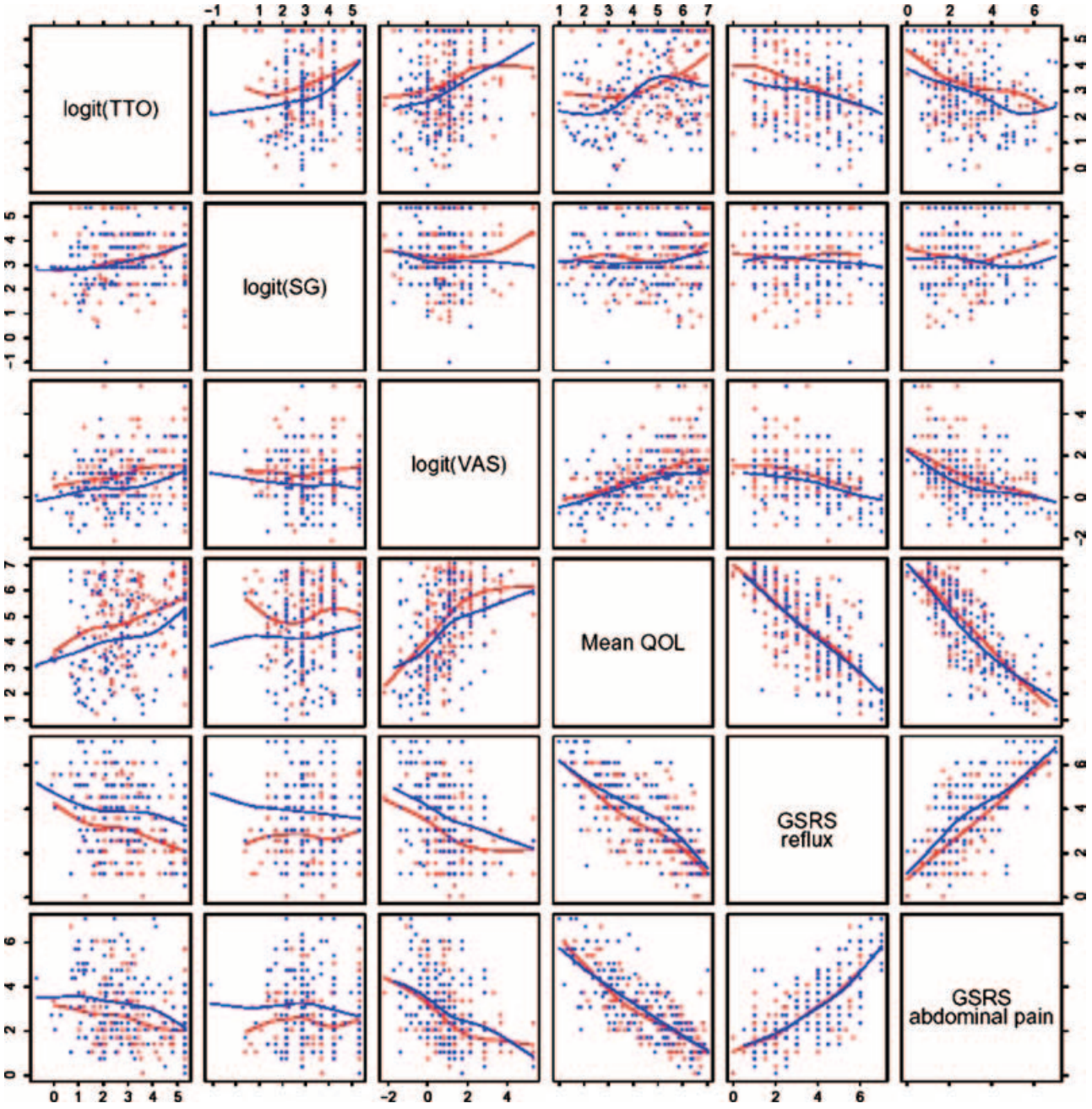

Figure 2. Scatterplot of the data. Logit transformations were applied to the time-tradeoff, standard-gamble, and visual-analog scores (with an adjustment to allow a logit of 100%). In each panel we have included a smooth regression curve, treating the vertical coordinate as the response, and the horizontal parameter as the predictor (separately for on (red) medication and off (blue) medication groups). There appears to be a weak correlation between the transformed time-tradeoff, standard-gamble, and visual-analog scores. The QOL score represents the mean score on the quality of life in reflux and dyspepsia survey. The GSRS reflux and abdominal pain scores were mean scores for each cluster of questions in the Gastrointestinal Symptom Rating Survey. The mean scores on the QOL and GSRS appear correlated, and all are correlated with logit (visual-analog score). Their correlations with the other two patient-derived utility measurements, the time-tradeoff, and standard-gamble values, are much weaker.

Since we studied a cohort of patients with daily reflux symptoms, we would anticipate that health-state utility values for patients with less frequent gastroesophageal reflux would probably be higher than the values that we obtained in the current study. In addition, the mean reflux symptom severity score on the GSRS was 3.9 (range 1–7) when patients were off of antisecretory therapy, representing that most of the patients with heartburn in this study had moderately severe symptoms. Patients with severe reflux symptoms might have been less likely to participate due to the requirement for temporary medication cessation. Future studies are needed in order to derive health-state utilities for patients with severe gastroesophageal reflux disease and for patients who are using intermittent or on-demand proton pump inhibitor

therapy. In addition, we did not collect data on the socioeconomic status or education level of our patients, which could have influenced our results.

We found that medication status at the time of the interview was an important factor for the time-tradeoff score, with scores off of antireflux therapy being significantly lower than on therapy ($p = 0.004$). While these results most likely reflect an increased severity of reflux symptoms off of antisecretory therapy, neither the patients nor the investigators were blinded to the medication status, and therefore the possibility of bias cannot be excluded. In addition, patients with comorbid conditions had time-tradeoff scores that were significantly lower when off of medications ($p = 0.002$), reflecting the importance of reflux symptoms even in the presence of other comorbid conditions.

There were no significant differences in standard-gamble scores when patients were on or off medications, and the time-tradeoff scores did not differ for patients with Barrett's esophagus or erosive esophagitis. In addition, the order in which patients were interviewed, either initially on or off medications, did not appear to influence the results, although the number of patients who were initially interviewed on medications was small. A potential weakness to our study was that we did not randomize half of our cohort to medication status at the time of interview. However, this change in study design would have been unlikely to change our primary results.

The values obtained in the current study using standard-gamble assessment were higher than those obtained from the time-tradeoff technique. Prior studies have shown a difference between values obtained, with the standard-gamble utilities exceeding the values obtained from time-tradeoff technique (27, 28). Prior studies have advocated that time-tradeoff utilities might reflect patient preferences for health more accurately than standard-gamble scores (29). Most individuals are risk averse and so they prefer to give up a stated amount of time, as in the time-tradeoff assessment, rather than to take a risk with uncertainty, as in standard gamble (30). Our results are consistent with this expected behavior. For example, a 50-yr-old male patient who trades away 1 month of his life in exchange for relief of heartburn symptoms would need to estimate a probability of death of 0.4% in the standard gamble to face an equivalent expected value in the number of years remaining.

An important strength of our methodology was the availability of endoscopic severity grading for all of the patients. The number of patients with erosive disease in our study was small (12%), reflecting the fact that most of the patients were using proton pump inhibitor therapy. Prior studies have suggested that patients with erosive disease do not have more severe symptoms when compared to patients with nonerosive disease (31, 32). The benefit of using health-state utilities compared to other measures of disease severity, such as the presence of esophageal erosions on upper endoscopy, is that utility values can more accurately reflect disease severity in addition to impact on health-related quality of life, particularly when patients are on proton pump inhibitor therapy and most likely will have normal upper endoscopic examinations.

A recent publication also used time-tradeoff, standard-gamble, and visual-analog scales in order to measure health-related quality of life in 1,011 heartburn patients in Germany and Sweden (33). In that study, the mean ($\pm$SD) scores were 0.88 ± 0.16 using time-tradeoff, 0.89 ± 0.14 with standard-gamble, and 0.69 ± 0.17 using a visual-analog scale. There are several important differences between the Kartman *et al.* analysis and our current study. First, patients in the European study were interviewed over the telephone and interviews were conducted using a hypothetical life expectancy of 10 yr. Patients' responses to complicated questions about life expectancy may be different when they are elicited by telephone compared to a personalized interview where patients can also process the information visually. Furthermore, individuals with gastroesophageal reflux disease span a broad age range and have correspondingly broad differences in life expectancy; our time-tradeoff method takes account of that by using the life expectancy as the time horizon rather than the fixed time horizon used by Kartman *et al.* Other differences between the two studies include confirmation of the diagnosis of heartburn in our patients with nonerosive disease by 24-hr pH monitoring, a requirement that symptoms worsen off of antireflux therapy and elicitation of utility values off of medications, stratification of results according to the presence of erosive disease or other comorbid conditions, and the inclusion of patients with Barrett's esophagus.

The visual-analog scale scores obtained were lower than the values for the time-tradeoff and standard-gamble values as is usually the case (34, 35). This result is expected since the visual-analog scale does not incorporate risk assessment. However, the visual-analog scores were able to provide face validity for the study results. As expected, the visual-analog scores were lower in patients with other comorbid conditions, and when patients were off medications for reflux symptoms.

We hypothesized that symptoms of reflux might be less important to patients with other comorbid conditions, and that the utility measurements might be higher in these patients. However, we did not find any significant difference between patients with or without comorbid conditions, or the presence of Barrett's esophagus, suggesting the importance of reflux symptoms to these patients. We found that duration of symptoms was an important predictor of the time-tradeoff score, and that a longer duration of illness correlated with a higher utility value. The likely reason for this result is that individuals tend to adjust to chronic illnesses over time, which might account for a higher score on a utility testing (36). Whether patients with Barrett's esophagus would have lower health-related quality of life scores if they were to be interviewed about the potential cancer risk associated with Barrett's esophagus is currently not known.

As expected, the scores on the disease-specific instruments improved with therapy for reflux, as has been shown in prior studies (37, 38). However, there was no correlation with Gastrointestinal Symptom Rating Survey or QOLRAD

Survey scores and utility values, illustrating that these disease-specific surveys may not accurately correlate with patients' assessment of health-related quality of life.

In summary, we have demonstrated that symptoms of gastroesophageal reflux disease adversely affect patients' health-related quality of life. The values defined in this analysis will be useful for future cost-effectiveness analyses in order to determine the most appropriate treatment for this chronic disorder. When obtaining quality of life information in patients with chronic gastroesophageal reflux disease, interviews should occur while patients are off medical therapy for heartburn, and using the time-tradeoff technique for utility assessment.

ACKNOWLEDGMENT

This study was supported by a Glaxo Institute for Digestive Health Healthcare Advancement Award and an AstraZeneca 2001 Quality Care Fund to Dr. Gerson. Dr. Goldstein's work was informed in part by NIH/NIA R01 AG15110. The views expressed are those of the authors and not necessarily those of the Department of Veterans Affairs.

Reprint requests and correspondence: Lauren B. Gerson, M.D., M.Sc., Division of Gastroenterology & Hepatology, Room A149, Stanford University Medical Center, Stanford, CA 94305-5202.

Received April 8, 2004; accepted August 3, 2004.

REFERENCES

1. Heartburn across America: A Gallup organization national survey. Princeton, NJ: Gallup Organization, 1988.
2. Nebel OT, Forbes MF, Castell DO. Symptomatic GE reflux: Incidence and precipitating factors. Dig Dis Sci 1976;21:953–6.
3. Isolauri J, Laippala P. Prevalence of symptoms suggestive of gastro-oesophageal reflux disease in an adult population. Ann Med 1995;27:67–70.
4. Revicki DA, Wood M, Maton PN, et al. The impact of gastroesophageal reflux disease on health-related quality of life. Am J Med 1998;104:252–8.
5. Johanson JF, Siddique R, Damiano AM, et al. Rabeprazole improves health-related quality of life in patients with erosive gastroesophageal reflux disease. Dig Dis Sci 2002;47:2574–8.
6. Stewart AL, Greenfield S, Hays RD, et al. Functional status and well-being of patients with chronic conditions. Results from the Medical Outcomes Study. JAMA 1989;262: 907–13.
7. Dupuy HJ. The psychological general well-being (PGWB) index. In: Wender NK, Mattson ME, Furber CD, Elinson J, eds. Assessment of quality of life in clinical trials of cardiovascular therapies. New York: Le Jacq, 1984;170–83.
8. Wiklund IK, Junghard O, Grace E, et al. Quality of life in reflux and dyspepsia patients. Psychometric documentation of a new disease-specific questionnaire (QOLRAD). Eur J Surg Suppl 1998;583:41–9.
9. Velanovich V, Vallance SR, Gusz JR, et al. Quality of life scale for gastroesophageal reflux disease. J Am Coll Surg 1996;183:217–24.
10. Gold MR, Patrick DL, Torrance GW, et al. Identifying and valuing outcomes. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996:82–134.
11. Goldstein MK, Tsevat J. Applying utility assessment at the "bedside." In: Chapman G, Sonnenberg F, eds. Decision making in health care: Theory, psychology, and applications. Cambridge: Cambridge Press; 2000:313–33.
12. Stalmeier PF, Goldstein MK, Holmes AM, et al. What should be reported in a methods section on utility assessment? Med Decis Making 2001;21:200–7.
13. Goldstein MK, Tsevat J. Assessing desirability of outcomes states for medical decision making and cost-effectiveness analyses. 2003; 1–14. Available from: Symptom Research: Methods and Opportunities at http://symptomresearch.nih.gov. National Institutes of Health, Bethesda, Maryland. Accessed 26 Jan 2004.
14. U.S. Decennial Life Tables Eliminating Certain Causes of Death. 1989–91: 18. Available from: National Center for Health Statistics at http://symptomresearch.nih.gov. Accessed 20 Jan 2004.
15. Nease RF, Tsai R, Hynes LM, et al. Automated utility assessment of global health. Qual Life Res 1996;5:175–82.
16. Lenert LA, Cher DJ, Goldstein MK, et al. The effect of search procedures on utility elicitations. Med Decis Making 1998;18:76–83.
17. Sumner W, Nease R, Littenberg B. U-titer: A utility assessment tool. Proc Annu Symp Comput Appl Med Care 1991;15:701–5.
18. Fryback DG, Lawrence WF Jr. Dollars may not buy as many QALYs as we think: A problem with defining quality-of-life adjustments. Med Decis Making 1997;17:276–84.
19. Harris RA, Nease RF Jr. The importance of patient preferences for comorbidities in cost-effectiveness analyses. J Health Econ 1997;16:113–9.
20. Mosteller F, Tukey J. Data analysis and regression. Boston, MA: Addison-Wesley, 1977.
21. Pinheiro J, Bates D. Mixed effects models in S and Splus. New York, NY: Springer-Verlag, 2000.
22. Nease RF Jr, Kneeland T, O'Connor GT, et al. Variation in patient utilities for outcomes of the management of chronic stable angina. Implications for clinical practice guidelines. Ischemic Heart Disease Patient Outcomes Research Team. JAMA 1995;273:1185–90.
23. Lewis EF, Johnson PA, Johnson W, et al. Preferences for quality of life or survival expressed by patients with heart failure. J Heart Lung Transplant 2001;20:1016–24.
24. Melsop KA, Boothroyd DB, Hlatky MA. Quality of life and time trade-off utility measures in patients with coronary artery disease. Am Heart J 2003;145:36–41.
25. Stavem K. Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease. Qual Life Res 1999;8:45–54.
26. Blumenschein K, Johannesson M. Relationship between quality of life instruments, health state utilities, and willingness to pay in patients with asthma. Ann Allergy Asthma Immunol 1998;80:189–94.
27. Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients' well-being and consequent effect on a cost-effectiveness analysis. J Clin Epidemiol 1992;45:505–12.
28. Lenert LA, Cher DJ, Goldstein MK, et al. The effect of search procedures on utility elicitations. Med Decis Making 1998;18:76–83.
29. Dolan P, Gudex C, Kind P, et al. Valuing health states: A comparison of methods. J Health Econ 1996;15:209–31.

30. Heath C, Tversky A. Preference and belief: Ambiguity and competence in choice under uncertainty. J Risk Uncertain 1991;4:5–28.
31. Gerson LB, Edson R, Lavori PW, et al. Use of a simple symptom questionnaire to predict Barrett's esophagus in patients with symptoms of gastroesophageal reflux. Am J Gastroenterol 2001;96:2005–12.
32. Johnson DA, Fennerty MB. Heartburn severity underestimates erosive esophagitis severity in elderly patients with gastroesophageal reflux disease. Gastroenterology 2004;126:660–4.
33. Kartman B, Gatz G, Johannesson M. Health state utilities in gastroesophageal reflux disease patients with heartburn: A study in Germany and Sweden. Med Decis Making 2004;24:40–52.
34. Hollingworth W, Deyo RA, Sullivan SD, et al. The practicality and validity of directly elicited and SF-36 derived health state preferences in patients with low back pain. Health Econ 2002;11:71–85.
35. Tosteson AN, Kneeland TS, Nease RF, et al. Automated current health time-trade-off assessments in women's health. Value Health 2002;5:98–105.
36. Tsevat J, Goldman L, Soukup JR, et al. Stability of time-tradeoff utilities in survivors of myocardial infarction. Med Decis Making 1993;13:161–5.
37. Havelund T, Lind T, Wiklund I, et al. Quality of life in patients with heartburn but without esophagitis: Effects of treatment with omeprazole. Am J Gastroenterol 1999;94:1782–9.
38. Talley NJ, Fullerton S, Junghard O, et al. Quality of life in patients with endoscopy-negative heartburn: Reliability and sensitivity of disease-specific instruments. Am J Gastroenterol 2001;96:1998–2004.