AHRQ Grant Final Progress Report

Title of Project: Preference Measurement for Multistate Health Profiles

Principal Investigator and Team Members:

- Julie A. Jacko, PhD, Associate Professor, Georgia Institute of Technology
- François Sainfort, PhD, Professor, Georgia Institute of Technology
- Thitima Kongnakorn, PhD, Post-Doctoral Researcher, Georgia Institute of Technology

Organization: Georgia Institute of Technology

Inclusive Dates of Project: 10/01/04 – 09/30/05

Federal Project Officer: Joan Metcalfe, Grants Management Officer

Acknowledgement of Agency Support: This project was supported by grant number R03-HS-14842 from the Agency for Healthcare Research and Quality

Grant Award Number: R03-HS-14842

1. Structured Abstract

Purpose: In measuring preference for multistate health profiles, past literature showed that the QALY (Quality-Adjusted Life Year) model, in which QALYs for each health state in the profiles are evaluated independently and all are combined to produce the overall QALYs, violates a key required assumption, additive independence. In this study, we aimed at investigating the relationship between two consecutive health states and exploring and understanding how the preference for a future health state is dependent upon the current health state.
Scope: The experiment was designed and performed with a group of healthy students in the context of general health described in EQ-5D system.
Methods: Two groups of subjects evaluated a set of hypothetical health scenarios described in EQ-5D system by using a visual analog scale (VAS) in one group, and time tradeoff (TTO) in the other group. The scenarios were constructed by varying three factors of interests: direction of change from current to future health states, amplitude of change, and current health state duration.
Results: Preference for a future health state was found to be strongly dependent on the level of current health state. The nature and extent of the impact of the current health state characteristics in assessing preference for a future health state also depend on the level of the future health state itself. The results have strong implications for cost-effectiveness studies.
Key Words: multistate health profiles, preference measurement, utility theory, QALY, additive independence

2. Purpose

The purpose of this study was to develop and test a new method that can better capture preferences for multistate health profiles. The motivation of the study came from the failure of the QALY (Quality-Adjusted Life Year) model in capturing preferences for multistate health profiles. As past literature shows, an existing technique that QALY uses in capturing preference for multistate health profiles, in which QALYs for each health state in the profiles are evaluated independently and all are combined to produce the overall QALYs, violates its required assumption, the additive independence[1-3]. Moreover, in preference assessment for multistate health profiles, the literature has shown that the preferences for entire health profiles are potentially impacted by many factors such as trend[4-6], rate of change[4, 7], peak and final intensity[7-10], spreading of outcomes[11], and timing of events[12, 13].

Based on these findings, this study proposed a novel approach to measure preference for multistate health profiles by assessing nested pairs of consecutive health states. The nature and extent of the interdependence between consecutive health states was expected to be revealed and used in assessing the overall profile. In addition, based upon the suggestion that people use reference points in evaluating the attractiveness of choices[14], a persuasive hypothesis that a health state will have an effect on one's preference for the following health state, which will have an effect on the preference of the next following state, and so on, was established. For example, a subject may evaluate the next health state, say health state X, as pleasant if his/her current health state is worse than health state X. On the other hand, the same subject may evaluate the same health state X as an unpleasant state if his/her current health state is better than health state X. Thus, decomposing the assessment of the entire health profile into a series of "conditional preference" assessments by looking at nested pairs of consecutive states was expected to better represent the overall preference score for the entire health profile than applying the commonly used QALY model since the interdependent relationships between health states would be taken into account. Thus, the specific aims of this study were as follows:

1. To explore if a preference score for future health state is dependent upon the level of current health state and identify which characteristics of the current health state impact the conditional preference scores for the future health state.

2. To test whether the proposed technique, which assesses "conditional preference scores" for multistate health states, can better predict preference scores for multistate health profiles than the current QALY model does.

3. Scope

A. Background

Violations of the Additive Independence Assumption

Several empirical research studies demonstrated the violation of the additive independence assumption in multistate health profile preference assessment. Richardson et al.[1] found that holistic preference scores of a three-state 16-year post-mastectomy health profile were significantly different from composite preference scores derived from the constituent health states. Moreover, Kupperman et al.[2] explored the predictability of a duration-weighted additive model in eight three-health state profiles in the context of prenatal diagnosis choices and found that the predictability of the duration-weighted additive model, as the conventional QALY model does, was not satisfied. Additionally, Spencer[3] found a significant difference between scores from holistic and constituent states elicitation in two out of the seven 10-year health profiles, each composed of up to three health states defined with the EQ-5D, which implied again that the additive independence assumption was violated.

Studies previously described clearly show the violation of the additive independence assumption. In addition, there are a number of studies that explored or identified profile characteristics that affect people's preferences for multistate profiles. These characteristics could lead to, and partially explain, the violation of the additive independence assumption. A review of the studies exploring such characteristics is as follows.

Rate of Change

Hsee and Abelson[15] performed experiments to find a relationship between satisfaction (utility) and rate of change of the outcomes in the contexts of gambling, class rank, and stock price. They found satisfaction to be positively related to actual outcome position and rate of change of the outcomes over time. Moreover, Chapman[4] found a significant effect of rate of change on the rating scores for 10 different profiles in the context of health and money. Subjects preferred gradually increasing or decreasing sequences to those with steep slopes. In a study of retrospective pain evaluation, the subjects experiencing heat stimuli on the forearms reported suffering significantly higher pain when the intensity steeply increased than when it gradually increased.[7]

Trend

Several empirical studies found a significant impact of trend of the overall profile (improvement versus decrement) on preferences.[4-8, 14, 16] For example, improving sequences were found to be strongly preferred to declining sequences in the domains of headache pain, athletic ability, facial acne and facial wrinkles.[5] In the context of health and money, improving sequences were preferred for short sequences (1 year) whereas for long sequences (lifetime), decreasing sequences were preferred for health but increasing ones for money.[4] Loewenstein and Sicherman[14] also showed that a majority of the subjects preferred an increasing sequence of wage profiles over a 5-year period rather than a declining one. In a very different domain, a majority of the subjects who reported having preference for a French restaurant over a Greek restaurant also reported, for sequences, preference for a dinner at a Greek restaurant first and at a French restaurant later, thus showing a preference for an improvement trend.[16]

Spreading of Outcomes

Decision makers prefer outcomes that spread across the time interval considered. For example, the majority of the subjects who were offered two free dinners preferred to distribute the two dinners across the time interval.[17] Moreover, in a study by Chapman[11] which involved scenarios including both gains and losses in the contexts of monetary outcomes (win a prize or pay a fine), dinner (pleasant or unpleasant dinner), and health-related events (a painful trip to the dentist or a pain-relieving trip to the chiropractor) in a four-week interval. In all scenarios, the majority of the subjects (70% to 92%) preferred to have the event on the first and the third weekends rather than on the first and the second weekends. About 65% to 91% of the subjects preferred to have the event happen on the second and the fourth weekends rather on the third and the fourth weekends.

Peak, Final Outcome, and Duration of the Profiles

A number of empirical studies have demonstrated that retrospective pain evaluation is influenced by the peak and final moment of the experience and not significantly impacted by the overall duration of the painful experience itself[8-10, 18, 19]. For example, Varey and Kahneman[9] asked 46 subjects to evaluate different discomfort profiles ranging from 15 to 35 minutes. They found that subjects' evaluations were significantly impacted by peak and final intensity but not from duration. The same phenomenon was also found in the retrospective evaluation of watching pleasant and unpleasant video clips[19], immersing hand in cold water[18], and in patients' retrospective evaluations of experiences in undergoing colonoscopy and lithotripsy[10].

Notice that the characteristics explored by the studies described previously were characteristics describing the overall pattern of the profile. No study has systematically explored the interdependence of preferences between each pair of two consecutive health states. In addition, based upon the suggestion that people use reference points in evaluating the attractiveness of choices[14], a persuasive hypothesis that a health state will have an effect on one's preference for the following health state, which will have an effect on the preference of the next following state, and so on, was established for this study. For example, a subject may evaluate the next health state, say health state X, as pleasant if his/her current health state is worse than health state X. On the other hand, the same subject may evaluate the same health state X as an unpleasant state if his/her current health state is better than health state X. Thus, this study aimed at exploring if a preference score for future health state is dependent upon the level of current (consecutively previous) health state and identify which characteristics of the current health state impact the preference score for the future health state.

B. Context

This experiment investigated preferences in the context of general health. Health states that were used in the experiment were described with the EQ-5D system[20], which presents a health state in terms of five dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. Each dimension is subdivided into three categories, which indicate whether the subject has no problem, a moderate problem, or an extreme problem. The EQ-5D system was chosen for this study since the description of a health state in those five dimensions was easy to understand and previous studies[21-23] had been performed to assist in the selection of appropriate health states for this study. Moreover, with such states, the subjects were not required to have knowledge of specific diseases.

C. Limitations of the study

One of the limitations involved the use of EQ-5D system to describe health states in the study. Although the description of the health states in five dimensions was easy to understand and did not require the subjects to have knowledge regarding diseases or any specific disease-related health conditions, the description of the five dimensions may not be specific enough and thus the subjects may have difficulty in imagining themselves in those health states and thus may have made additional unknown assumptions regarding the health states being considered and evaluated. Take, for example, the fourth dimension concerning pain or discomfort, subjects needed to make an assumption about the types of pain or discomfort or specific symptoms that are the results of that pain or discomfort. However, the EQ-5D system remains an acceptable and widely used tool in health outcome measurement today.

Another limitation concerns one of the assessment techniques used to capture preferences in this study: visual analog scale (VAS). While the VAS is attractive because it is simple, easy to administer, and the subjects can perform the task by themselves, the validity of the VAS approach is still questionable. VAS scores represent the strength of the preference under certainty. Thus, VAS is not a utility measurement technique and thus cannot fully represent individuals' preferences under risk and uncertainty. Moreover, some biases may be associated with the use of VAS. For example, the VAS score for a health state also depends on the number of better or worse health states presented previously or at the same time. If many better health states are shown before or simultaneously with the health state assessed, then the score assigned might be undervalued. On the other hand, if several worse health states are shown before or at the same time with the health state assessed, its score might be overvalued. However, even with known issues, VAS has been acceptable and is widely used in similar research and studies for health outcome measurement and health economics. Using the VAS technique, while not a gold standard of utility measurement, allowed us to perform the study and revealed important relationships.

D. Participants

The subjects in this study were undergraduate or graduate students of the School of Industrial and Systems Engineering at the Georgia Institute of Technology. This subject pool represented a young and relatively healthy population. One rationale behind using a healthy population in the study was that in economic evaluation of health care, the valuations of health used in cost effectiveness analysis are societal and thus include healthy subjects as opposed to patients with actual experience with diseased or morbid states[24]. While the sample used in this study was certainly not representative of the general population, it represented one segment of the general population. More importantly, the purpose of this exploratory research was to investigate, develop and test a new assessment approach, not to generate utility assessments that can be used in actual cost-effectiveness studies.

4. Methods

A. Study Design

For ease of presentation, the experimental design was divided into two phases corresponding to the two aims of the study. The methods are then presented that describe the flow of the experiment for each subject. Each subject was exposed to both phases in sequence.

A.1 Research Design for Aim 1 - Phase 1

The first specific aim was to explore the relationship between two consecutive health states (current and future health states) by investigating factors of the current health state that potentially impact the conditional preference score for the future health state. As a result of the literature review, the following three factors were of interest.

Factor#1: Direction of change from current health state

The direction of change from current health state to future health state contributes to the perceived preference for the future health state. If the future health state is increased from the current one, it will be perceived as a gain, otherwise as a loss. The same future health state X might be assigned a smaller preference score when perceived as a loss (Figure 1A) than when perceived as a gain (Figure 1B).

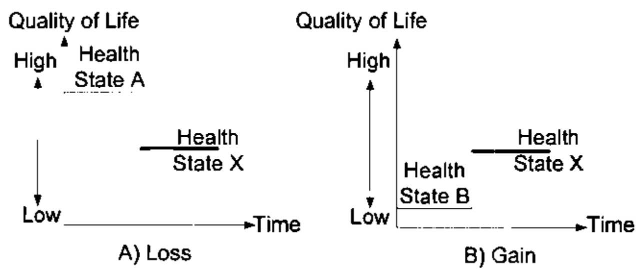

Figure 1: Direction of Change
(Same future health state X perceived as a loss in situation A and as a gain in situation B)

Factor #2: Amplitude of change from current health state to future health state

Amplitude indicates the absolute change between two health states (e.g. current health state and future health state, in this case). Amplitude of change from current health state to future health state was suspected to affect the conditional preference score of the future health state because people may use their current health state as a reference point when evaluating their future health states. Different amplitudes may yield different preference scores for the same future health state. For example, if an individual is asked to value a future health state X, as shown in Figure 2, will he/she give the same score if the health state that he/she is currently in is different (health state A vs. health state B)?

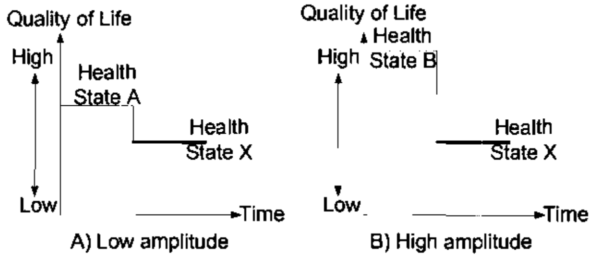

Figure 2: Amplitude of Change

(Same future health state X but with small loss amplitude in situation A and large loss amplitude in situation B)

<u>*Factor#3: Current health state duration*</u>

The duration that an individual has been in his/her current health state might also impact the conditional evaluation of the future health state. For example, if an individual is asked to value his/her future health state, health state X, shown in Figure 3, will he/she give the same score for health state X if he/she has been staying in his/her current health state for a short duration (A) versus a long duration (B)?

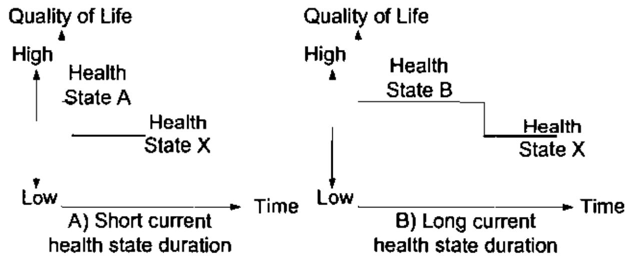

Figure 3: Current Health State Duration

(Same future health state X but with short current state duration in A and long current health state duration in B)

Thus, in phase 1, a full 2^3 factorial design (three factors with two levels each) for a total of 8 scenarios was employed. Each scenario consisted of two health states: current and future health state. The level of future health state was fixed across all 8 scenarios. However, the current health state was varied based on the level of the three factors designed. Figure 4 graphically illustrates the 8 scenarios. The two levels for each factor that were used in the experiment are as follows:

1. <u>Direction of change</u> from the current health state to the future health state: increase vs. decrease
2. <u>Amplitude of change</u> (on a utility scale from 0 to 1): large (0.30) vs. small (0.15).
3. <u>Duration of current health state</u>: long (10 years) vs. short (1 year)

In order to get better and more reliable results, replications of the experiment with different levels of the future health states were employed. With the limitation that health state weights ranged from 0 to 1 and the required amplitude range (0.30 for large amplitude) in the experiment, the highest possible level of the future health state examined was 0.70. Indeed, a current health state better than the future health state by an amount of 0.30 (large amplitude) leaded to a value of 0.70+0.30=1.00, the maximum on the scale. Similarly, the lowest level of the future health state had to be greater than 0.30 since a current health state worse than the future health state by a large amplitude of 0.30 leaded to a value of 0.30-0.30=0.00, the minimum on the scale. With these restrictions, a total of three levels for the future health state were employed, taking on values of 0.70, 0.55, and 0.40, which will be referred to as Design A, B, and C from this point. Figure 9 shows the levels of current and future health states required in the experiments (different durations are not shown in the figure). Thus, a total of 3x8=24 two-state profiles, or scenarios, were used to form three sets of 8 profiles, each corresponding to a full factorial design and corresponding to one of three levels of the future health state.

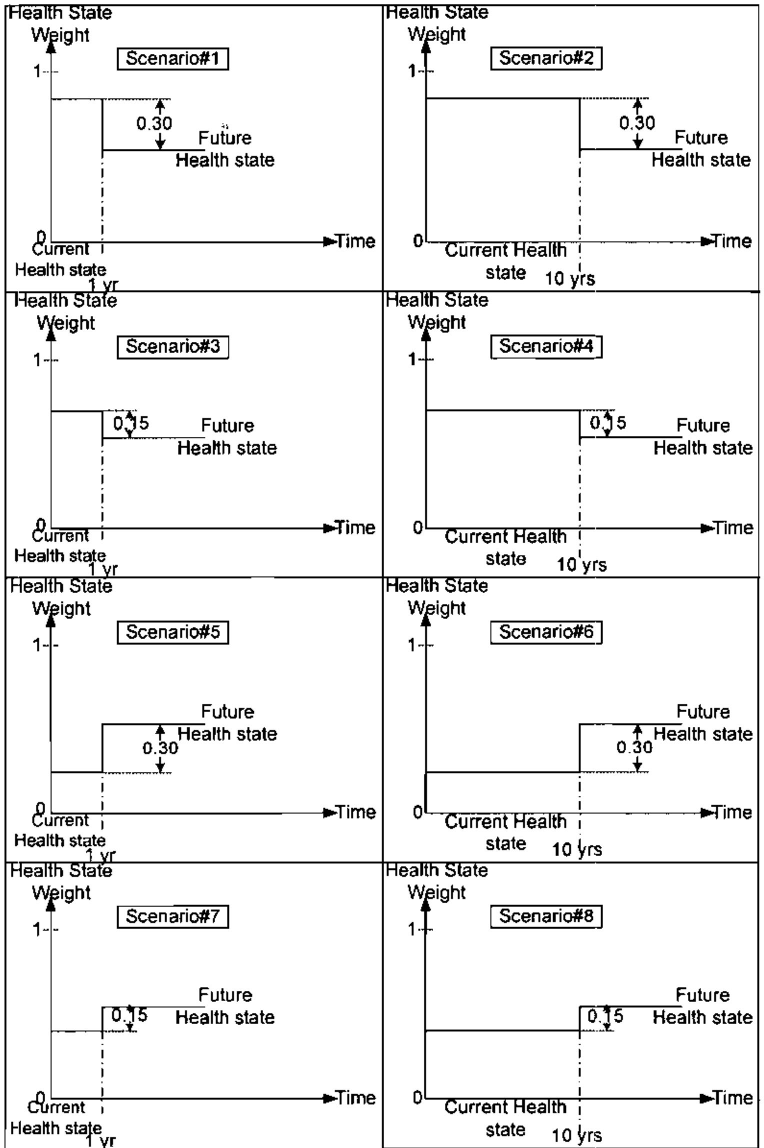

Figure 4: Graphical representation of a set of 8 scenarios in the 2^3 factorial design

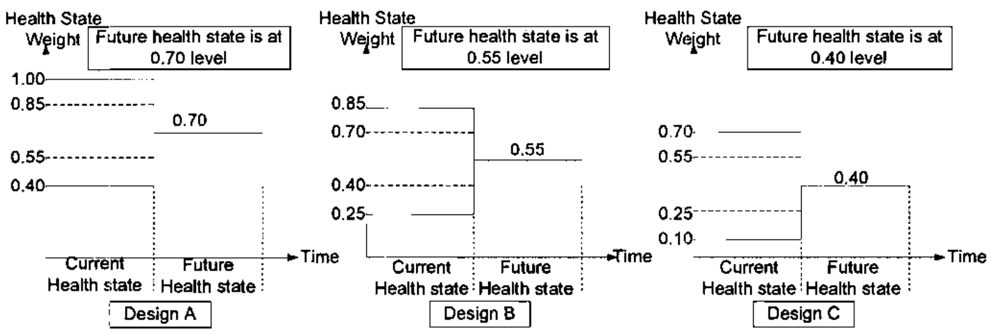

Figure 5: Graphical illustrations of three full designs corresponding to 3 different levels of the future health state *(Note: not shown in the figure, each current health state will take on two different durations, short vs. long)*

A.2 Research Design for Aim 2 - Phase 2

The experimental design in Phase 2 was set up to accomplish the second aim of the study: to test if the proposed technique, the "conditional preference assessment", can better predict preference scores for an entire health profile than the current unconditional health states assessment method commonly used in the QALY model. Thus, the only factor of interest was the health profile evaluation technique itself. Three types of evaluation techniques were as follows:

1. Holistic preference score assessment

The holistic assessment was obtained by presenting the subjects with the entire multistate health profile and asking them to evaluate the entire profile.

2. Unconditional preference score assessment

The unconditional assessment, corresponding to the conventional QALY model, was obtained by presenting the subjects with each health state in the multistate health profile and asking them to evaluate each health state independently.

3. Conditional preference score assessment

The conditional assessment, the proposed method for this study, was obtained by assessing the individual preference score for the first state followed by the conditional preference score of each future state given the current state for all nested pairs of consecutive states. For example, for a health profile that consists of 4 health states, health state 1 is evaluated independently. State 2 is evaluated given state 1, state 3 is evaluated given state 2 and finally state 4 is evaluated given state 3.

For Phase 2 experiment, a total of 10 health profiles were constructed by varying different levels of the three factors previously introduced (direction, amplitude, and duration). Figure 6 shows the 10 health profiles that were constructed. Regarding the conditional preference scores, most of the conditional preference scores for each pair of two consecutive health states in 10 health profile could be obtained from Phase1 experiment. However, there were 7 additional scenarios that were needed for Phase 2 but were not obtained for Phase 1 which were also added to the experiment.

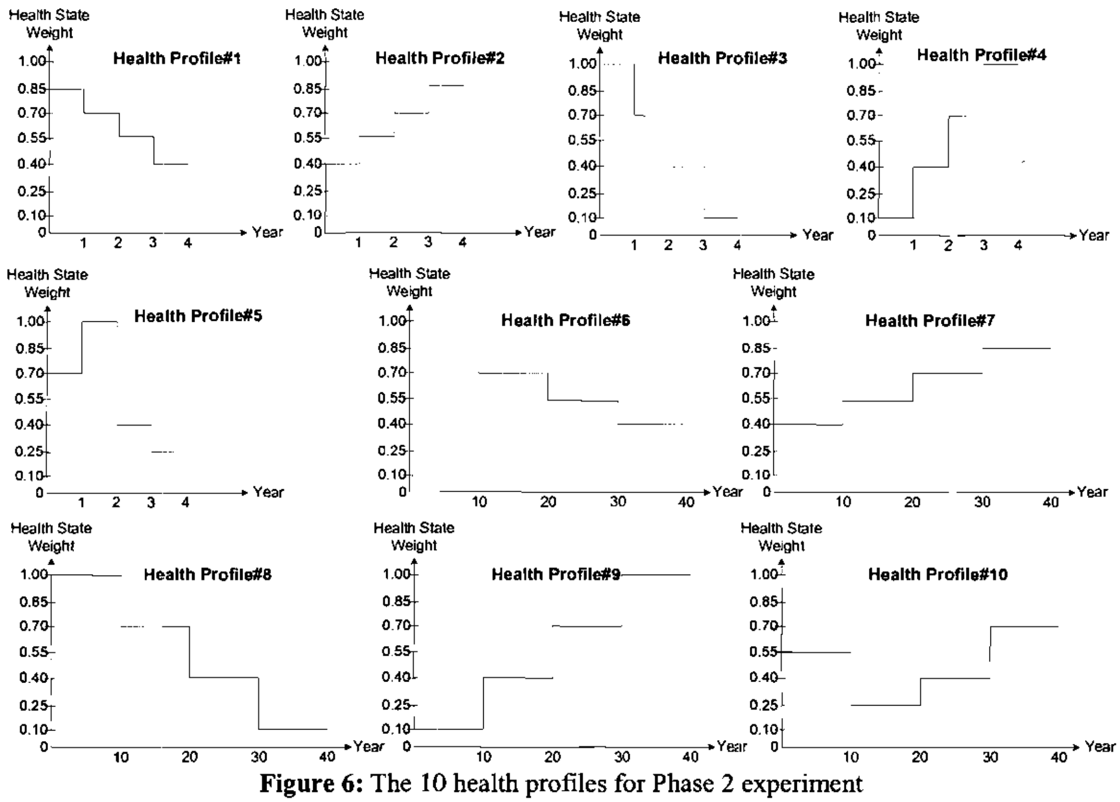

Figure 6: The 10 health profiles for Phase 2 experiment

A.3 Hypotheses:

1. Preference score for future health state is dependent upon the level of the current health state.
2. The proposed decomposition method in which several conditional preference score assessments are performed and integrated to produce a preference score for the entire health profile will predict the holistic preference of the entire health profile better than the decomposition technique that uses unconditional preference score assessments (the conventional QALY model).

A.4 Elicitation Techniques

Two elicitation techniques were used for both phase 1 and phase 2, thus leading to two separate studies for both aims: visual analog scale (VAS) and time-tradeoff (TTO). In the VAS assessment, the subjects were instructed to indicate how they feel about the health states by marking a 0 to 100 scale with 2 extremes: best imaginable state (score 100) and worse imaginable state (score 0). With TTO, an individual was presented with the task of deciding how much time they would be willing to trade off to be in a better health state versus a poorer one[25]. The subject was asked to trade off time in the healthier state versus the less healthy state, essentially by further decreasing or increasing the time in the healthy state until he/she was indifferent.

B. Development of the Experiment

Each experiment started by choosing and designing the experimental scenarios that would be presented to the subjects. Health states that were used in the experiment were described with the EQ-5D system[20].

Because different people may have different judgments regarding the same health states, designing the experiment by using the same set of health states across all subjects may result in uncontrollable perceived amplitude, one of the factors that was investigated and thus needed to be controlled. Hence a within-subject standardization process was used to choose appropriate scenarios for each subject. Therefore, at the beginning of the experiment, each subject was presented with 20 descriptions of health states and asked to evaluate each of them independently. From this step, unconditional preference scores obtained would be used to construct health scenarios for Phase 2 experiment. Due to the complexity and large time-consumed of TTO method, only 10 health states were presented and evaluated. The same set of 20 (10 for TTO) health states were used for all subjects. Then, health states that had preference scores that were closest to the scores needed for the experiment (0.85, 0.70, 0.55, 0.40, 0.25, and 0.10 (for level 1.00, perfect health was employed)) were selected for use in the remainder of the experiment for that particular subject. Thus, while all subjects started with the same 20 states (10 for TTO), the subset chosen for the rest of the experiment may differ from subject to subject but were corresponding to the health states and profiles demanded by the designs described earlier. There were three major parts for each of the two experiments (VAS and TTO) as follows:

1. Unconditional preference assessments with 20 health states (10 for TTO)

For the VAS study, this part was done manually by using 20 index cards. Each card had a description of an individual health state. First the subjects were asked to rank order the 20 health states from the best to the worst health states. Then, unconditional preference assessments were obtained for each of the 20 health states on VAS. The scores from the assessments were then entered into the computer program for selection of health states to be used in parts 2 and 3. In this part, basic demographic information was also collected from the subjects (gender, ethnicity, age, major field of study, degree worked on, experience with major health issue, general health). Regarding the TTO study, each subject performed the TTO assessment for each of the 10 health states on a computer, which was programmed particularly for this study.

2. Conditional preference assessments

Based on the 20 (10 for TTO) unconditional preference scores from the assessments in Part 1, the computer program devised for this study selected the required health states and generated health scenarios for conditional preference assessments. A series of health scenarios were presented in a random order. Each scenario displayed the hypothetical current health state along with the current health state duration followed by a future health state. The subject was asked to assess the future health state in the scenario given the current scenario. For VAS study, there were a total of 31 health scenarios (24 scenarios necessary for Phase 1 plus an additional 7 scenarios necessary for profiles introduced in Phase 2). However, regarding the TTO study, since the data for VAS study was collected before the TTO study. Many subjects in VAS study reported having difficulty in imagining the difference between having been in the current health state in 1 year and in 10 years. Thus, they reported that duration factor was not taken into account much when they performed the assessment. Therefore, in the TTO study, only two factors (direction and amplitude) were investigated in order to minimize the experimental task due to the complexity of TTO. Three replications of a 2^2 factorial design were employed which yielded a total number of 12 health scenarios. However, an additional 2 health scenarios were needed for Phase 2 experiment. Thus, the total number of health scenarios for conditional preference assessment for TTO study was 14 scenarios.

3. *Holistic preference assessments*

For the VAS study, the 10 health profiles (Figure 6) were generated by the computer program based on the 20 unconditional preference scores in the first part. Those 10 health profiles were divided into two groups based on the profiles duration: five 4-year profiles, and five 40-year profiles. The scenarios of the 4-year profiles were framed in terms of the conditions during an illness lasting 4 years, while the 40-year profiles were framed as lifetime profiles. Before the holistic preference assessments started, the subjects performed paired-comparisons for all possible combinations of the five profiles within each group with the aim of generating the ranking among the five profiles. Then, the subjects were asked to assess each health profile. Regarding the TTO study, since the TTO method is not appropriate for evaluating temporary or short-term health states[26, 27], only lifetime profiles were included in the experiment.

The computer program developed for this study employed the Visual Basic programming language. The program was self-explanatory so that the subjects could perform the tasks by themselves. Each subject took approximately 45 minutes to complete the experiment.

C. Data Collection

Two independent experiments were conducted. Data for VAS study was collected from 100 students in the fall semester of year 2004. Regarding the TTO study, the data was collected from 85 students in the summer semester of year 2005. The study was advertised through electronic mail by targeting undergraduate and graduate students of the School of Industrial and Systems Engineering at the Georgia Institute of Technology (a total population of over 1,500 students). Those interested in participating in the study made an appointment with the investigator of the study and came to perform the study at the Laboratory for Human-Computer Interaction and Health Care Informatics co-directed by Drs. Jacko and Sainfort.

The experiment started by giving a short explanation of the study and asking if the student agreed to participate in the study. A signed consent form was obtained for each participant. Then, each subject performed the experiment on the computer. Each subject received $20 as compensation.

5. Results

A. Data Analyses

Data from VAS and TTO study was separately analyzed. First, outlier analysis was performed using box plot analysis. Data identified as outliers were removed which leaded to the remaining 92 data for the VAS study and 78 data for the TTO study. Among the 92 data in the VAS study, 36 (39%) were female students and 56 (61%) were male students. Age ranged from 18 to 33 years, with an average of 22.36 years. Regarding the ethnicity, 44 (48%) were white, 28 (30%) were Asian or Pacific Islander, 10 (11%) were Hispanic, and 6 (7%) were black. The majority (83 (90%)) of them majored in Industrial Engineering and 72 (78%) were undergraduate students. Concerning major health issues, 13 (14%) experienced a major health issue themselves, 48 (52%) experienced one in their families, and 40 (43%) experienced one in someone else close to them. None of them indicated having poor health on the experimental day. Concerning the TTO study, 25 (32%) were female students and 53 (68%) were male students. Age ranged from 18 to 34 years, with an average of 22.41 years. The participants consisted of, 39 (50%) whites, 30 (38%) Asians or Pacific Islanders, 3 (4%) Hispanics, 3 (4%) blacks, and 3 (4%) others. Sixty-seven (86%) of them majored in Industrial Engineering and 55 (70%) were undergraduate students. Concerning major health issues, 8 (10%) experienced a major health issue themselves, 45 (58%) experienced one in their families, and 37 (47%) experienced one in someone else close to them. Moreover, none of them indicated having poor health.

Unconditional Preference Scores

Unconditional preference scores were obtained when the preference score for each health state were assessed independently of other health states. Descriptive statistics of the unconditional preference scores for the 6 health states (at the designed levels 0.85, 0.70, 0.55, 0.40, 0.25, and 0.10) required by the experiment from VAS and TTO study were presented in Table 1. It can be seen that health states selected to construct health scenarios for Phase 2 experiment had average scores ranging from 0.057 to 0.842 for VAS study while the scores from TTO was higher (ranged from 0.305 to 0.939). The average amplitude between each consecutive level of the required health state ranged from 0.15 to 0.16 and from 0.09 to 0.15 for VAS and TTO study respectively.

Table 1: Unconditional preference score for each level of health state selected for the experiment

	VAS	**TTO**
Health State	**Mean (S. D.)**	**Mean (S.D.)**
Health State Level 0.85	0.842 (0.029)	0.939 (0.081)
Health State Level 0.70	0.686 (0.032)	0.846 (0.126)
Health State Level 0.55	0.522 (0.035)	0.733 (0.150)
Health State Level 0.40	0.365 (0.047)	0.588 (0.214)
Health State Level 0.25	0.208 (0.042)	0.441 (0.249)
Health State Level 0.10	0.057 (0.036)	0.305 (0.256)

Conditional Preference Scores (CPS)

Conditional preference scores were obtained by presenting each subject with a set of health scenarios. A health scenario consisted of a description of a current health state along with the duration that the subject was asked to imagine himself/herself currently experiencing, followed by a description of a future health state that the subject would experience starting the next day. Table 2 shows descriptive statistics of CPS obtained from all subjects for Design A, B, and C (future health state level 0.70, 0.55, and 0.40, respectively) for all scenarios (8 scenarios per design for VAS study, 4 scenarios per design for TTO study).

Table 2: Conditional Preference Score for Each Scenario in Design A, B, and C

Factors			**Design A**		**Design B**		**Design C**	
			VAS	**TTO**	**VAS**	**TTO**	**VAS**	**TTO**
Direction	**Amplitude**	**Duration**	**Mean (S.D.)**	**Mean (S.D.)**	**Mean (S.D.)**	**Mean (S.D.)**	**Mean (S.D.)**	**Mean (S.D.)**
Decrease	Large	1 year	0.410(0.179)	0.815(0.167)	0.545(0.130)	0.663(0.247)	0.404(0.133)	0.545(0.257)
Decrease	Large	10 years	0.663(0.168)		0.539(0.138)		0.422(0.142)	
Decrease	Small	1 year	0.674(0.121)	0.810(0.195)	0.535(0.128)	0.683(0.244)	0.412(0.124)	0.564(0.276)
Decrease	Small	10 years	0.670(0.137)		0.534(0.131)		0.414(0.136)	
Increase	Large	1 year	0.645(0.147)	0.780(0.221)	0.480(0.157)	0.672(0.257)	0.442(0.210)	0.572(0.279)
Increase	Large	10 years	0.644(0.162)		0.518(0.160)		0.492(0.206)	
Increase	Small	1 year	0.644(0.127)	0.812(0.200)	0.534(0.122)	0.668(0.242)	0.394(0.144)	0.511(0.270)
Increase	Small	10 years	0.644(0.144)		0.538(0.130)		0.408(0.153)	

In order to test the first hypothesis, General Linear Model – Repeated Measures (GLM-RM) was used to analyze the potential effects of interested factors (direction, amplitude, and duration) on the conditional preference scores. This technique was used as opposed to between-groups analysis of variance (ANOVA) since the same subjects participated in all conditions of the experiment. SPSS 11.5 for Windows was used as the analysis tool for running GLM-RM.

GLM-RM was performed by treating duration, direction, and amplitude (only direction and amplitude for the TTO study) as repeated factors, while conditional preference score was a dependent variable. For the VAS study, the analysis was performed on the original data since it was normally distributed. However, for the TTO study, since normality assumption was not met on the original data, the analysis was performed on the transformed data (1-square root(1-TTO)). Three sets of GLM-RM were executed separately for each design (Design A, B, and C). Summary of the main effects, 2-way interaction, and 3-way interaction effects for VAS study and TTO study are summarized in Table 3.

Table 3: Results from GLM-RM for each design

	Design A		**Design B**		**Design C**	
	VAS	**TTO**	**VAS**	**TTO**	**VAS**	**TTO**
Effects	**F (p-value)**	**F (p-value)**	**F (p-value)**	**F (p-value)**	**F (p-value)**	**F (p-value)**
Duration	47.24 (<0.001)*	N/A	1.34 (0.250)	N/A	6.42 (0.013)*	N/A
Direction	8.11 (0.005)*	0.84 (0.361)	2.87 (0.094)	0.00 (0.962)	2.65 (0.107)	0.28 (0.599)
Amplitude	58.31 (<0.001)*	1.83 (0.180)	2.96 (0.089)	0.21 (0.648)	14.48 (<0.001)*	2.38 (0.127)
Duration x Direction	62.04 (<0.001)*	N/A	3.04 (0.085)	N/A	1.58 (0.212)	N/A
Duration x Amplitude	75.38 (<0.001)*	N/A	1.23 (0.271)	N/A	2.72 (0.102)	N/A
Direction x Amplitude	46.91 (<0.001)*	1.28 (0.262)	4.21 (0.043)*	1.35 (0.249)	9.47 (0.003)*	8.53 (0.005)*
Duration x Direction x Amplitude	68.63 (<0.001)*	N/A	2.06 (0.155)	N/A	0.43 (0.511)	N/A

*Significant different at alpha = 0.05

As expected, it was found that characteristics of the current health state had an effect on preference judgment for the FHS. However, the results obtained from VAS and TTO studies were different. For TTO study, only Design C showed a significant effect of direction and amplitude interaction. While for VAS study, Design A had the most significant effects when compared to the results from the other two designs. Moreover, unexpectedly, the nature of the effects varied across different levels of the FHS, indicating that in addition to the fact that preference for a future health state depends on current health state, the nature and extent of the impact of the current health state's characteristics in assessing preference for a future health state (i.e. direction of change, amplitude of change, and current health state duration) also depend on the level of the future health state itself. The significant effects will be explained by each Design in the section below.

Design A (Future health state level 0.70)

For the VAS study, all main effects, 2-way interaction, and 3-way interaction effects were significant for Design A. Since the 3-way interaction effect was significant, all the main effects and 2-way interactions can be ignored. As shown in Figure 7, it can be seen that direction and amplitude effects interacted only when current health state had a 1-year duration. When current health state had 1-year duration and future health state had small amplitude of change from current health state, the mean of CPS for a future health state that was decreased from the CHS was slightly higher than the mean CPS for a future health state that was increased from the current health state (0.674 versus 0.644). However, when current health state had 1-year duration and future health state had large amplitude of change from current health state, changing the direction of change from *decreasing* to *increasing* significantly increased CPS by an average of 0.235. Regarding the TTO study, none of the effects were significant.

Design B (Future health state level 0.55)

For the VAS study, only 2-way interaction effect between direction and amplitude was significant. Figure 8 shows that when future health state was decreased from current health state, the CPS for the future health state that had large amplitude of change from current health state was slightly higher, with an average of 0.007, than the CPS for the future health state that had small amplitude of change. On the other hand, when future health state was increased from current health state, CPS for future health state that had large amplitude was lower than the CPS for future health state that had small amplitude of change from current health state by an average of 0.037 score. Regarding the TTO study, none of the effects were significant.

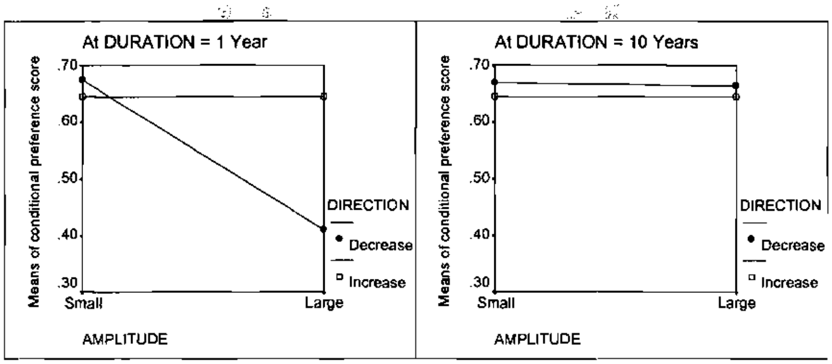

Figure 7: Graphs represent three-way interaction effect on conditional preference score for Design A from VAS study (future health state level = 0.70)

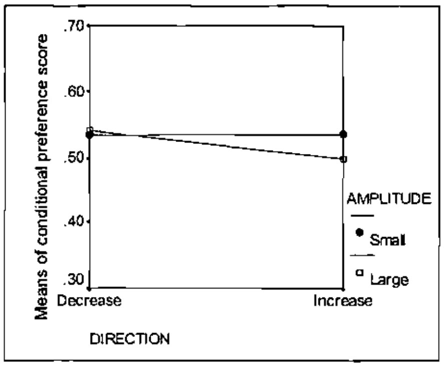

Figure 8: Interaction plots represent 2-way interaction effect between direction and amplitude for Design B (future health state level = 0.55)

Design C (Future health state level 0.40)

For the VAS study, 2-way interaction effect between direction and amplitude, main effects of duration and amplitude were significant. Regarding the 2-way interaction, Figure 9 shows that for a future health state that was decreased from current health state, CPS for future health state that corresponded to a large amplitude of change from current health state was approximately equal to CPS for future health state that had small amplitude of change from current health state. However, for a future health state that was increased from current health state, varying the amplitude of change from small to large significantly increased CPS by an average of 0.066 score. Since this 2-way interaction was significant, the main effect of amplitude can be ignored. Regarding the main effect of duration, from Figure 9, for future health state at level 0.40, CPS when current health state had a 1-year duration was significantly lower than CPS when current health state had a 10-year duration; with an average of 0.021 score of difference.

Regarding the results from TTO study, the 2-way interaction effect between direction and amplitude had significant effect on CPS. See Figure 10, for a future health state that was decreased from current health state, CPS for future health state that corresponded to a large amplitude of change from current health state was lower than CPS for future health state that had small amplitude of change from current health state by an average of 0.019 score. However, for a future health state that was increased from current health state, varying the amplitude of change from small to large significantly increased CPS by an average of 0.061 score.

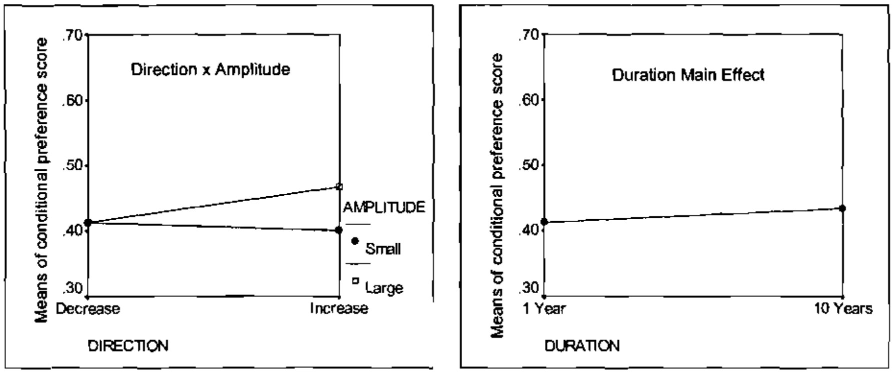

Figure 9: An interaction plot represent 2-way interaction effect between direction and amplitude and a plot represents main effect of duration on CPS for Design C (future health state level = 0.40)

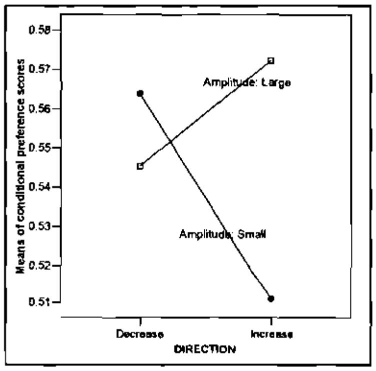

Figure 10: An interaction plot represent 2-way interaction effect between direction and amplitude for Design C (future health state level = 0.40) (TTO study)

Holistic Preference Scores

For each subject, a holistic preference score was obtained for each health profile by asking a subject to rate each health profile on a VAS or assess an indifferent point for the TTO study. Each subject generated 10 holistic preference scores for a total of 10 health profiles for the VAS study and 4 scores for 4 health profiles for the TTO study (only lifetime profiles were assessed since TTO assessment is not appropriate for short-term health state assessment). Table 4 shows descriptive statistics of holistic preference scores for each health profile.

Table 4: Holistic Preference Scores for Each Health Profile

	VAS	TTO
Holistic Preference Scores	**Mean (S.D.)**	**Mean (S.D.)**
Health Profile#1 (gradually decrease, 4 yr)	0.591 (0.188)	N/A
Health Profile#2 (gradually increase, 4 yr)	0.778 (0.136)	N/A
Health Profile#3 (steeply decrease, 4 yr)	0.206 (0.174)	N/A
Health Profile#4 (steeply increase, 4 yr)	0.572 (0.249)	N/A
Health Profile#5 (no pattern, 4 yr)	0.421 (0.194)	N/A
Health Profile#6 (gradually decrease, lifetime)	0.756 (0.157)	0.690 (0.179)
Health Profile#7 (gradually increase, lifetime)	0.655 (0.172)	0.718 (0.216)
Health Profile#8 (steeply decrease, lifetime)	0.466 (0.238)	0.571 (0.198)
Health Profile#9 (steeply increase, lifetime)	0.271 (0.193)	0.617 (0.309)
Health Profile#10 (no pattern, lifetime)	0.481 (0.211)	N/A

Notice that for the 4-year health profiles (profiles#1 to 5), increasing profiles had higher averageVAS scores than decreasing profiles. On the other hand, for the lifetime health profiles (profiles #6 to 10), decreasing profiles have higher average VAS scores than increasing profiles. However, for the TTO study, in which only the lifetime profile preference scores were assessed, the results were consistent with the results from VAS study for the 4-year health profiles, not the lifetime health profiles. The increasing profiles had higher preference scores than decreasing profiles. Another interesting point is that the scores obtained from TTO were not much fluctuated (ranged from 0.571 to 0.718) as compared to those from VAS study (ranged from 0.271 to 0.756).

In order to test the second hypothesis, *duration-weighted scores (DWS)* were calculated and compared across the three assessment techniques: conditional, unconditional, and holistic preference assessments. Descriptive statistics for DWS for each health profile are shown in Table 5. Then, Friedman tests with Wilcoxon Signed Rank tests for post-hoc tests (non-parametric) and GLM-RM with Bonferroni adjustment for post-hoc tests (parametric) were performed. Note that DWSs from TTO study were normally distributed while those from VAS study were not. Thus, GLM-RM analyses were used for the data from TTO study, and Friedman tests were used for the data from VAS study. Results from Friedman tests and GLM-RM are shown in Table 6. It can be seen that for the VAS study, there was no significant difference in DWS among the three assessment techniques for health profiles 1, 4, and 10. For health profiles 2, 6, and 7, DWS-holistic was found to be significantly higher than DWS-conditional and DWS-unconditional. The opposite finding was found for health profiles 3, 5, 8, and 9, in which DWS-holistic was found to be significantly lower than DWS-conditional and DWS-unconditional. Regarding the comparison between DWS-conditional and DWS-unconditional, profiles 2, 3, and 7 have significantly higher DWS-unconditional than DWS-conditional, while for profiles 5, 6, and 9, DWS-unconditional is significantly lower than DWS-conditional. However, these results are not consistent with the findings from TTO study (profiles 6 to 9), DWS-holistic is found to be significantly lower than DWS-unconditional in all profiles, and lower than DWS-conditional in profiles 6, 7, and 8. Moreover, DWS-unconditional is found to be significantly higher than DWS-conditional in profiles 6 and 7.

Table 5: DWS for each health profile for each assessment technique

	VAS Mean (S.D.)			TTO Mean (S.D.)		
	Conditional	**Unconditional**	**Holistic**	**Conditional**	**Unconditional**	**Holistic**
Profile #1	2.465 (0.23)	2.421 (0.09)	2.364 (0.75)	N/A		
Profile #2	2.338 (0.22)	2.421 (0.09)	3.110 (0.54)	N/A		
Profile #3	1.910 (0.30)	2.113 (0.09)	0.823 (0.70)	N/A		
Profile #4	2.143 (0.26)	2.113 (0.09)	2.286 (0.99)	N/A		
Profile #5	2.618 (0.22)	2.263 (0.09)	1.683 (0.78)	N/A		
Profile #6	24.716 (2.12)	24.214 (0.86)	30.240 (6.27)	29.958 (5.89)	31.263 (0.09)	27.609 (7.17)
Profile #7	23.537 (2.44)	24.214 (0.86)	26.214 (6.88)	29.671 (6.11)	31.263 (4.35)	28.712 (8.62)
Profile #8	21.738 (2.40)	21.127 (0.86)	18.643 (9.51)	27.349 (5.70)	27.500 (4.78)	22.840 (7.90)
Profile #9	22.051 (2.58)	21.127 (0.86)	10.844 (7.71)	26.696 (5.68)	27.500 (4.78)	24.667 (12.34)
Profile #10	18.496 (2.48)	17.866 (1.09)	19.226 (8.45)	N/A		

Another approach to look at the performance of the proposed assessment technique, conditional preference assessment, was to look at the preference between pairs of the health profiles that the conventional method, unconditional preference assessment, failed to predict. Those health profiles were health profiles that consisted of the exact same health states but were in different sequences: health profiles# 1 and 2, 3 and 4, 6 and 7, and 8 and 9. For those pair of profiles, unconditional preference assessment would generate equal preference scores.

Analysis at this stage was performed in order to explore if conditional preference assessment can capture the preferences between pairs of health profiles that unconditional preference assessment fails to capture. The choice predicted from conditional preference score assessment was compared to the actual choice from paired comparison. As described earlier, in Phase 2 experiment, the subjects performed paired-comparisons for all possible combinations of the five profiles within each group (4-year profiles group and lifetime profiles group). In this analysis, for each subject, the preferred profiles assessed from conditional preference assessment were identified by the higher DWS-conditional between each health profile pairs of interest. Then the preferred profiles identified were compared to the preferred profiles obtained from the paired comparisons. Table 7 shows the percentage of subjects in which the preferred profiles from conditional preference assessment match those from the paired comparisons.

Table 6: Results from Friedman Tests and GLM-RM
(H = DWS-holistic, UC = DWS-unconditional, and C = DWS-conditional)

Health Profiles		Profile 1	Profile 2	Profile 3	Profile 4	Profile 5	Profile 6	Profile 7	Profile 8	Profile 9	Profile 10
VAS (Friedman Tests)	χ^2	0.161	80.276	113.954	1.103	108.621	57.379	11.793	8.759	79.966	1.402
	p-value	0.923	<0.001*	<0.001*	0.576	<0.001*	<0.001*	0.003*	0.013*	<0.001*	0.496
	Post Hoc Tests		H > UC H > C UC > C	H < UC H < C UC > C		H < UC H < C UC < C	H > UC H > C UC < C	H > UC H > C UC > C	H < UC H < C UC < C	H < UC H < C UC < C	
TTO (GLM-RM)	F	N/A	N/A	N/A	N/A	N/A	9.638	9.411	17.767	4.484	N/A
	p-value						<0.001*	<0.001*	<0.001*	0.015*	
	Post Hoc Tests						H < UC H < C UC > C	H < UC H < C UC > C	H < UC H < C	H < UC	

*Significant different at alpha = 0.05

Table 7: Percentage of subjects in which the preferred profiles from conditional preference assessment match those from the paired comparisons

Comparison between health profiles#	**Percentage of matched cases**	
	VAS	**TTO**
1 and 2	36%	N/A
3 and 4	70%	N/A
6 and 7	56%	56%
8 and 9	44%	53%

The results showed that conditional preference assessment can predict the preferences between health profiles that the unconditional technique failed to predict up to 70% match with the prediction from paired comparison which is a much more reliable technique in predicting which profile is preferred to another.

B. Conclusion and Discussion

In this study, a conditional preference assessment technique, a new method to capture preferences for multistate health profiles was developed and tested. The experiments were designed and conducted to investigate its validity. The hypotheses were formulated and examined in two phases of the study. In the first phase of the study, the potential effect of the relationship between current health state and future health state on preference judgments of the future health state was explored. Subjects were presented with different hypothetical health scenarios. Each scenario was composed of two different health states: a current health state and a future health state. For each scenario, a conditional preference score for future health state was elicited using a direct rating through a visual analog scale (VAS) in one study, and a time-tradeoff (TTO) method in another study. Various scenarios were created by varying direction of change between health states, amplitude of change, duration of the current health state, and level of the future health state. The results from the VAS and TTO study were analyzed separately for each level of the future health state. The results can be summarized by each design as follows:

- Design A – All main and interaction effects were significant in the VAS study. However, none of the effects were significant in the TTO study.
- Design B – Only 2-way interaction effect between direction and amplitude was significant in the VAS study. None of the effects were significant in the TTO study.
- Design C – Main effects of duration and amplitude were significant in the VAS study. Moreover, the 2-way interaction between direction and amplitude were significant in both VAS and TTO studies.

As expected, it was found that characteristics of the current health state had an effect on preference judgments for future health states. However, unexpectedly, the nature of the effects varied across different levels of the future health state, indicating that in addition to the fact that preference for a future health state depended on current health state, the nature and extent of the impact of the current health state characteristics in assessing preference for a future health state (i.e. direction of change, amplitude of change, and current health state duration) also depended on the level of the future health state itself.

The results showed many significant effects on the conditional preference scores in the VAS study while there was only one significant effect in the TTO study. The fewer significant results in the TTO study could be because of the possible bias associated with the TTO technique: the loss aversion. Loss aversion refers to the finding that people are more sensitive to losses than to gains[28]. Because a person was asked to trade off life years (a loss) for a perfect health (a gain), TTO scores tended to be overestimated due to the person's loss aversion. As can be seen from the results of this study, the scores assessed from TTO were clustering around the upper end of the utility scale. Those scores then were hard to be detected as significant different among one another.

With respect to the results from the VAS study, all effects were significant for health scenarios at a higher level of future health state (Design A – future health state level at 0.70). While only one and three out of the seven effects for health scenarios at a medium level of future health state (Design B – future health state level at 0.55) and a lower level of future health state (Design C – future health state level at 0.40), respectively, were significant. These results imply that preferences for future health state were more sensitive to the current health state as the future health state was better.

In order to better explain and understand the findings regarding the interdependence between current health state and future health state on the preference scores, additional analysis, the Wilcoxon Signed Ranks test, was performed in order to test the difference between the mean conditional preference score and the mean actual future health state score (when assessed unconditional) assigned by the subjects, for each scenario. Table 8 summarizes the discrepancies and the test results. (Note that the positive value of discrepancy indicates that the actual future health state score was higher than its conditional preference score, and vice versa).

Table 8: Mean of the discrepancy between the actual future health state score and the conditional preference score along with the p-values from Wilcoxon Signed Ranks test

<table>
<tr><th colspan="3" rowspan="2">Factors</th><th colspan="2">Design A</th><th colspan="2">Design B</th><th colspan="2">Design C</th></tr>
<tr><th>VAS</th><th>TTO</th><th>VAS</th><th>TTO</th><th>VAS</th><th>TTO</th></tr>
<tr><th>Direction</th><th>Amplitude</th><th>Duration</th><th>Mean (p-value)</th><th>Mean (p-value)</th><th>Mean (p-value)</th><th>Mean (p-value)</th><th>Mean (p-value)</th><th>Mean (p-value)</th></tr>
<tr><td>Decrease</td><td>Large</td><td>1 year</td><td>0.276(<0.001)*</td><td rowspan="2">0.031(0.088)</td><td>-0.023(0.068)</td><td rowspan="2">0.063(0.011)*</td><td>-0.039(0.009)*</td><td rowspan="2">0.033(0.091)</td></tr>
<tr><td>Decrease</td><td>Large</td><td>10 years</td><td>0.023(0.346)</td><td>-0.017(0.206)</td><td>-0.057(0.001)*</td></tr>
<tr><td>Decrease</td><td>Small</td><td>1 year</td><td>0.013(0.531)</td><td rowspan="2">0.036(0.095)</td><td>-0.013(0.365)</td><td rowspan="2">0.046(0.056)</td><td>-0.047(<0.001)*</td><td rowspan="2">0.025(0.501)</td></tr>
<tr><td>Decrease</td><td>Small</td><td>10 years</td><td>0.016(0.793)</td><td>-0.012(0.215)</td><td>-0.049(0.001)*</td></tr>
<tr><td>Increase</td><td>Large</td><td>1 year</td><td>0.041(0.005)*</td><td rowspan="2">0.066(0.033)*</td><td>0.042(0.006)*</td><td rowspan="2">0.060(0.056)</td><td>-0.077(0.003)*</td><td rowspan="2">0.010(0.686)</td></tr>
<tr><td>Increase</td><td>Large</td><td>10 years</td><td>0.042(0.014)*</td><td>0.004(0.572)</td><td>-0.127(<0.012)*</td></tr>
<tr><td>Increase</td><td>Small</td><td>1 year</td><td>0.042(0.002)*</td><td rowspan="2">0.034(0.105)</td><td>-0.012(0.388)</td><td rowspan="2">0.061(0.023)*</td><td>-0.029(0.254)</td><td rowspan="2">0.072(0.005)*</td></tr>
<tr><td>Increase</td><td>Small</td><td>10 years</td><td>0.042(0.004)*</td><td>-0.016(0.433)</td><td>-0.043(0.037)*</td></tr>
</table>

* Significant difference at alpha = 0.05

The results from Table 8 show that conditional preference scores were significantly different from the actual future health state scores, indicating that current health state affects the judgment of the future health state, in 13 out of 24 scenarios for VAS study and 4 out of 12 scenarios for TTO study. Again, not many significant results in TTO were found and this could due to the loss aversion associated with TTO previously described. An interesting finding found in the VAS study is that the mean conditional preference score in Design A was always *lower* than the actual future health state score (positive values of the discrepancies). On the other hand, in Design C, conditional preference score was always *higher* than the actual future health state score (negative values of the discrepancies). In Design B, there was virtually no significant difference between conditional preference score and actual future state score. This overall finding is important: independently of direction of change, amplitude of change and duration of current health state, scoring a future health state with information about the current health state affected the future health state score downward if the future health state was high and upward if the future health state was low. These findings signify that the pattern of the relationship between current health state and the future health state on the judgment of the future health state depended upon the level of the future health state itself. Thus the results strongly support that preferences for future health states strongly and systematically depend on where the current health state is. Moreover, the pattern of the interdependence also depends upon the level of future health state itself. Another interesting finding was found in the results from TTO study. The actual future health state score was always *higher* than its conditional preference score in all designs. This finding may indicate that with the TTO technique, conditioning the future health state on the level of current health state affected the preference score downward regardless of the relative level of current health state to future health state.

In the second phase of the study, the aim was to investigate whether the proposed decomposition technique (conditional preference assessment) can better predict preference score for a full health profile than the conventional decomposition technique (unconditional preference assessment). Subjects were presented with different hypothetical health profiles and were asked to assess holistic preference scores for each health profile by providing a direct rating through VAS for or assessing a utility score by TTO. Each health profile was composed of four different health states. Duration-weighted scores were calculated for scores obtained from conditional, and unconditional preference assessments and were compared to holistic scores and to each other.

Unfortunately, the results from the analyses of both VAS and TTO studies did not show great promise that the proposed decomposition technique that used conditional preference score assessments predicted holistic preference scores for health profiles significantly better than the decomposition technique that used unconditional preference score assessments. However, when looking at the holistic preference assessment itself, it should be recognized that the assessment tasks themselves can be problematic especially when the number of health states in the profile becomes large. It is still questionable if holistic preference scores assessed, regardless of whichever strategies individuals use, are fully reliable in terms of representing their preferences regarding the hypothetical health profiles. Several studies as well as this study have proposed other decomposition techniques in order to aid with the difficulties that arise in formulating holistic preference score assessment.

Although the analyses did not show that conditional preference scores predicted holistic preference scores better than unconditional preference scores, some of the results indicatde that conditional preference assessment can predict preferred choices between pairs of health profiles better than unconditional preference assessment. This is important since, while individuals have difficulties in evaluating health profiles (i.e., producing a holistic score), they fare much better in selecting preferred profiles in forced paired comparisons, as designed in this study. In the case that unconditional preference assessment cannot predict trend of preference due to the additive independence assumption, for example, profiles that are composed of the exact same components of health states but different in sequences, unconditional preference assessment would automatically produce identical duration weighted scores across the profiles, indicating that those profiles are equally preferred. In contrast, conditional preference assessments in which the interdependence relationship between a pair of health states is taken into account, were generally, although not all the time, indicative of the actual direction of preference in this case. Thus, from the perspective of paired comparisons across health profiles, the conditional method performed better than the unconditional method. Moreover, the results from Phase 1 experiment strongly support the fact that an interdependence relationship between current health state and future health state exists and matters in making preference assessments. Thus, it is essential for further research to perform future studies in order to develop a method that is able to integrate the interdependence relationship between health states into the preference assessment technique in the case of multistate health profiles. The simple method proposed here, while a good first step did not completely live up to expectations in terms of solving the problem.

C. Implications

Concerning the implications for other related fields, the results from this study showed that current health state and its characteristics had an effect on individuals' judgments regarding their future health states. Thus, practitioners and clinicians, who want to capture patients' preferences regarding their health states in order to make use of them in medical decision making, should definitely ensure that every future health state considered is conditioned on the patients' current health state. Regarding the implications from a societal perspective, specifically cost effectiveness analyses, this study suggests that the effectiveness of interventions designed to improve health for people with very low current health state might be underestimated. The results from this study show that people with very low current health state value their future health state higher than its actual level. Thus, with the assessment from the existing decomposition technique, the effectiveness of the interventions may be undervalued. This implies that some interventions that are not currently seen as cost effective might in fact be cost effective.

D. Directions for Future Research

The results from this study confirm that there is a relationship between two consecutive health states that has an effect on the judgment of the future health state. However, this study did not successfully develop a new measurement method that was able to integrate all the significant relationships uncovered into a new calculation of preference scores for evaluating complex health profiles. Thus, future research would be important to be performed in at least two general directions. First, future research is in need to further explore the nature and extent of the relationship between health states such as exploring complex relationships between more than two consecutive health states. And second, performing similar experimental studies with different subject populations, such as using actual patients (with experienced health states) instead of young healthy subjects (with hypothetical health scenarios) would be another important direction for future research in multistate health profile preference assessment.

6. List of Publications and Products

Kongnakorn T, Sainfort F, Jacko JA. Preference measurement for multistate health profiles. Poster to be presented at the Society for Medical Decision Making's 27th Annual Meeting 2005 Oct 15-18; San Francisco, CA.

References

1. Richardson J, Hall J, Salkeld G. The measurement of utility in multiphase health states. International Journal of Technology Assessment in Health Care 1996;12:151-62.
2. Kuppermann M, Shiboski S, Feeny D, Elkin EP, Washington E. Can Preference scores for Discrete States Be Used to Derive. Preference Scores for an Entire Path of Events? Med Decis Making 1997;17:42-55.
3. Spencer A. A test of the QALY model when health varies over time. Soc. Sci. Med. 2003;57:1697-1706.

4. Chapman GB. Expectations and Preferences for Sequences of Health and Money. Organizational Behavior and Human Decision Processes 1996;67(1):59-75.
5. Chapman GB. Preferences for Improving and Declining Sequences of Health Outcomes. Journal of Behavioral Decision Making 2000;13:203-218.
6. Krabbe PFM, Bonsel GJ. Sequence Effects; Health Profiles, and the QALY Model: In Search of Realistic Modeling. Med Decis Making 1998;18:178-88.
7. Ariely D. Combining Experiences Over Time: The Effects of Duration, Intensity Changes and On-Line Measurements on Retrospective Pain Evaluations. Journal of Behavioral Decision Making 1998;11:19-45.
8. Ariely D, Zauberman G. On the making of an experience: The effects of breaking and combining experiences on their overall evaluation. Journal of Behavioral Decision Making 2000;13:219-232.
9. Varey C, Kahneman D. Experiences extended across time: evaluation of moments and episodes. Journal of Behavioral Decision Making 1992;5:169-85.
10. Redelmeier DA, Kahneman D. Patients' memories of painful medical treatments: real-time and retrospective evaluaitons of two minimally invasive procedures. Pain 1996;66:3-8.
11. Chapman GB. Sooner or later: The psychology of intertemporal choice. The psychology of learning and motivation 1998;38:83-113.
12. Bleichrodt H, Johannesson M. Time preference for health: A test of stationarity versus decreasing timing aversion. Journal of Mathematical Psychology 2001;45:265-282.
13. Chapman GB. Temporal discounting and utility for health and money. Journal of Experimental Psychology: Learning, Memory and Cognition 1996;22(3):771-791.
14. Loewenstein G, Sicherman N. Do workers prefer increasing wage profiles? Journal of Labor Economics 1991;9:67-84.
15. Hsee CK, Abelson RP. Velocity Relation: Satisfaction as a Function of the First Derivative of Outcome Over Time. Journal of Personality and Social Psychology 1991;60(3):341-7.
16. Loewenstein GF, Prelec D. Negative time preference. American Economic Review 1991;81:347-52.
17. Loewenstein GF, Prelec D. Preferences for Sequences of Outcomes. Psychological Review 1993;100(1):91-108.
18. Kahneman D, Fredrickson BL, Schreibner CA, Redelmeier DA. When more pain is preferred to less: adding a better end. Psychological Science 1993;4:401-5.
19. Fredrickson BL, Kahneman D. Duration neglect in retrospective evaluations of affective episodes. Journal of Personality and Social Psychology 1993;65(1):45-55.
20. The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199-208.
21.Kind P, Dolan P, Gudex C, Willaims A. Variations in population health status: results from a United Kingdom national questionnaire survey. BMJ 1998;316:736-741.
22. Dolan P. Modeling valuations for EuroQol health states. Medical Care 1997;35:1095-1108.
23. Busschbach JJV, McDonnell J, Essink-Bot ML, van Hout BA. Estimating parametric relationships between health description and health valuation with an application to the EuroQol EQ-5D. Journal of Health Economics 1999;18:551-571.
24. Hadorn D. The role of public values insetting health care priorities. Soc. Sci. Med. 1991;32:773-781.
25. Torrance G, Thomas WH SD. A utility maximization model for evaluation of health care programs. Health Service Research 1972;7:118-133.
26. Cook J, Richardson J, Street A. A cost utility analysis of treatment options for gallstone disease: methodological issues and results. Health Economics 1994;3:157-168.
27. Swan J, Fryback D, Lawrence W, Sainfort F, Hagenauer M, Heisey D. A time-tradeoff method for cost-effectiveness models applied to radiology. Med Decis Making 2000;20:79-88.
28. Tversky A, Kahneman D. Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain 1992;5:297-323.