Preferences for Health Outcomes and Cost-Utility Analysis

George W. Torrance, PhD

Abstract

Economic evaluation of health programs consists of the comparative analysis of alternative courses of action in terms of both costs and consequences. The five analytic techniques are cost-consequence analysis, cost-minimization analysis, cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis. Although all techniques have the same objective of informing decision making in health programs, they come from different theoretic backgrounds and relate differently to the discipline of economics.

Cost-utility analysis formally incorporates the measured preferences of individuals for the health outcome consequences of the alternative programs. The individuals may be actual patients who are experiencing or have experienced the outcomes, or they may be a representative sample of the community, many of whom may someday face the outcomes. The health outcomes, at the most general level, consist of changes in the quantity and quality of life; that is, changes in mortality and morbidity. Changes in quantity of life are measured with mortality; changes in quality of life are measured with health-related quality-of-life instruments.

Utilities represent a particular approach to the measurement of health-related quality of life that is founded on a well specified theory and provides an interval scale metric. Changes in quantity of life, as measured in years, can be combined with changes in quality of life, as measured in utilities, to determine the number of quality-adjusted life years gained by a particular health program. This can be compared with the incremental cost of the program to determine the cost per quality-adjusted life-year gained.

Utilities may be measured directly on patients or other respondents by means of techniques such as visual analog scaling, standard gamble, or time trade-off. Utilities may be determined indirectly by means of a preference-weighted multi-attribute health status classification system such as the health utilities index. The health utilities index is actually a complete system for use in studies. It consists of questionnaires in various formats and languages, scoring manuals, and descriptive health status classification systems. The health utilities index is useful in clinical studies and in population health surveys, as well as in cost-utility analyses.

(Am J Man Care 1997;3(Suppl):S8-S20)

Cost-utility analysis is one of the analytic approaches used in the economic evaluation of health programs. A special feature of cost-utility analysis is that it incorporates the expressed preferences of relevant individuals for the outcomes produced or averted by the program. There are two main groups of individuals generally seen as relevant in measuring these pref-

From the Centre for Health Economics and Policy Analysis, McMaster University; Centre for Evaluation of Medicines, St. Joseph's Hospital; Department of Clinical Epidemiology and Biostatistics, McMaster University; and Innovus Research, Inc., Burlington, Ontario, Canada

Address correspondence to: George W. Torrance, PhD; Health Sciences Centre 3H1C, McMaster University, 1200 Main St. West, Hamilton, Ontario, Canada L8N 3Z5.

erences—patients and the general public. In some sense, however, both groups are patients, either current or future. Members of the general public are potential future patients for most of the programs or health outcomes under consideration in these analyses. This paper provides an introductory overview to the concepts and techniques of cost-utility analysis with special emphasis on the methods of measuring and incorporating the preferences.

Economic Evaluation

Purpose. Economic evaluation is a general umbrella term that refers to a collection of analytic techniques all aimed at facilitating decision making regarding the overall merit of healthcare programs considering both costs and consequences. The methods essentially are designed to help decision makers discriminate between programs that represent good "value for money" from those that do not. The techniques are not always referred to as *economic evaluation* but sometimes go under the rubric of *technology assessment.* The important point is that the word *economic* in economic evaluation does not necessarily refer to the discipline of economics. It can simply refer to the casual use of the adjective *economic* to reflect the concept of costs and value for money. Indeed, cost-utility analysis did not grow out of the discipline of economics. Rather, it was developed as a branch of cost-effectiveness analysis, founded on decision theory and operations research.[1-4] Recently, there have been attempts to link cost-utility analysis formally with the discipline of economics.[5,6] This link, however, is not essential to cost-utility analysis, and the presence or absence of this link does not change in any way how cost-utility analysis is practiced.

Economic evaluation is defined as the comparative analysis of alternative courses of action, considering both their costs and consequences, to inform decision making. There are a number of important points in the definition. The analysis must always be comparative. A healthcare program cannot be analyzed in isolation. There is always an alternative program, and often that alternative program is simply whatever would have happened in the absence of the program under consideration. Because these analyses are always comparative, one is always considering the incremental costs and the incremental consequences. The techniques of economic evaluation can be used to study large national programs that cost hundreds of millions of dollars. They can equally well be used to compare small choices that one may have in delivering specific treatments. The techniques are quite general and are applicable both at the micro level and at the macro level. Economic evaluation requires the explicit consideration of both the costs and the consequences of a program. Like everything else in life, we are concerned with both what we pay and what we get. It would make no sense to base decisions only on costs or only on consequences. It is important to underscore that the purpose of economic evaluation is to inform and facilitate decision making, not to replace decision making. That is, economic evaluation is a decision aid, not a decision-making tool. It is not meant to replace common sense, judgment, and careful deliberation, all of which are still required and are crucial.

Like everything else in life, we are concerned with both what we pay and what we get. It would make no sense to base decisions only on costs or only on consequences. It is important to underscore that the purpose of economic evaluation is to inform and facilitate decision making, not to replace decision making.

Techniques of Economic Evaluation. There are five analytic techniques of economic evaluation: cost-consequence analysis, cost-minimization analysis,

cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis. In cost-consequence analysis the comparative data on costs and consequences are tabulated but are not combined or analyzed. If there are different types of consequences, they are simply listed separately. The approach is similar to that taken by *Consumer Reports*, the well-known US consumer products testing magazine. A cost-consequence analysis would seldom be seen as sufficient on its own, but can often be a used as a helpful display of the data before one or more of the other techniques are used.

Cost-minimization analysis is appropriate if the alternatives being compared have identical consequences. That is, if the effectiveness of the two interventions is equal, including side effects and adverse events, the decision can be made strictly on the basis of costs.

In cost-effectiveness analysis the consequences are measured in their natural units, such as blood pressure reduction in millimeters of mercury, cases found, patients improved, lives saved, and life-years gained. The cost-to-effectiveness ratio is expressed in terms of the cost-per-unit gain of the primary effectiveness measure.

The distinguishing feature of cost-utility analysis is that the consequences not only are counted (as in cost-effectiveness analysis) but also are valued. That is, the consequences considered more important, as based on expressed preferences of patients or the general public, are given greater weight. The most widely used approach to cost-utility analysis is to measure the consequences in terms of quality-adjusted life years (QALYs) gained.

In cost-benefit analysis the consequences are valued by converting them into an equivalent monetary amount. Willingness to pay is the technique used to monetize the consequences. The cost-benefit approach then is used to compare costs in monetary units with consequences in monetary units to determine whether the program provides an overall benefit.

Foundations of Economic Evaluation. To better understand the various techniques of economic evaluation it is helpful to review the theoretic or disciplinary foundations of each. Cost-consequence analysis is designed to be a practical decision aid. It is based on no particular theory but is simply an organized way to summarize and present the key data to the decision maker. Cost-minimization analysis, cost-effectiveness analysis, and cost-utility analysis all were originally founded on decision sciences and operations research, particularly optimization theory.[1,2] They all are related to each other, and they were all developed to facilitate decision making. Cost-utility analysis has attracted additional debate regarding its theoretic foundations, and two possible foundations have been identified. The original foundation, based on optimization theory and the maximization of QALYs, has been termed the *extra welfare foundation.*[7] In this approach, the maximization of QALYs is considered an appropriate social goal for the healthcare system. As an alternative, Garber et al developed a welfare economics foundation for cost-utility analysis.[5,6] Cost-benefit analysis is founded squarely on welfare economics. Although this does give a strong theoretic basis for cost-benefit analysis, some argue that important aspects of this basis are problematic for use in healthcare.[8] The main concern is that consequences are monetized on the basis of willingness to pay, as constrained by the ability to pay for these consequences, by the patients, potential patients, or the general public. The issue is simply that many people reject the notion that ability to pay for consequences should factor into decision making in healthcare programs. The concern is that such an approach favors the programs and diseases of the affluent over those of the poor.[8]

Health-Related Quality of Life

Definition and Rationale. Quality of life is a broad concept that includes many aspects of life in addition to health; for

example, wealth, freedom, democracy, and cleanliness of the environment all contribute to one's overall quality of life. Health-related quality of life refers to those aspects of quality of life that are tied to health.

The overall goal of healthcare is to make patients feel better and live longer. Physiologic measures (eg, sedimentation rate, forced expiratory volume, serum creatinine) do not necessarily correlate with feeling better. It is important to measure health-related quality of life. Moreover, many treatments that have a beneficial effect on the disease also produce undesirable side effects for some patients. Health-related quality of life measures allow one to capture both the good and the bad effects of treatment and to combine them in one composite measure.

Consumer sovereignty is an important concept in economics. The notion is that the consumer is the best judge of his or her own welfare. In terms of health this means that the consumer is the best judge of which outcomes are more desirable, and by how much. Health-related quality of life provides a method to use consumer sovereignty in the evaluation of health programs. It is now widely agreed that health-related quality of life is an important measure to incorporate in studies.[9-11]

Taxonomy of Measures of Health-Related Quality of Life. Health-related quality of life instruments can be classified into three groups: specific instruments, generic health profiles, and preference-based measures.[12,13] Specific instruments include those targeted at specific diseases, such as the functional living index—cancer[14] and the Western Ontario–McMaster osteoarthritis index[15]; specific populations, such as the care and resource evaluation tool for the elderly[16]; and specific functions, such as visual function measured with the activities of daily vision scale[17,18] or knee function measured with the rating system developed by the Knee Society.[19]

Generic health profiles are applicable to a wide range of patients and diseases. They provide scores on a number of dimensions and typically are not aggregated into an overall summary score. The data are displayed as a profile across the dimensions, hence the name for this group of instruments. Three well-known instruments in this category are the SF-36,[20] the sickness impact profile,[21] and the Nottingham health profile.[22]

Preference-based measures provide a single summary score for health-related quality of life. The score is based on the measured preference of individuals for the health-related quality of life associated with the consequences of the program. There are two basic types of these instruments: direct and indirect. The direct instruments are used to measure the preferences of individuals with direct techniques such as the visual analog scale, time trade-off, standard gamble, willingness to pay, or special instruments such as the McRheum.[23-26] Indirect instruments are much simpler to use. A patient's health status is classified into a system that provides a preference-based score. Three well-known systems in this category are the quality of well-being,[27,28] the health utilities index (HUI),[29-31] and the EuroQol-5D.[32,33]

Requirements for Use in Cost-Effectiveness or Cost-Utility Analysis. The preference-based measures of health-related quality of life potentially can be used in cost-utility analysis to provide the quality-adjustment factors needed to determine quality-adjusted life years gained. However, to be appropriate for this use, the measures must meet the following requirements. They must be preference-based and interval-scaled, and perfect health and death must be on the scale.[11] A preference-based measure ensures that consequences that are more preferred will receive a greater weight in the analysis than those that are less preferred. An interval scale is required to support the calculations and statistical tests required. Perfect health as well as death must be on the scale because these outcomes occur in programs and must be incorporated. All the instru-

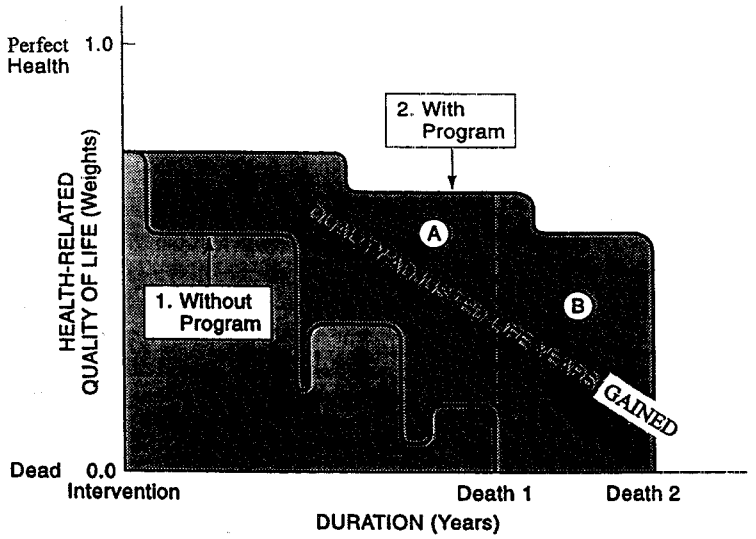
ments listed earlier as preference-based measures meet these requirements. It is important to note, however, that generic health profiles, the best known of which is the SF-36, even if they produce a single summary score, do not meet the requirements for use in cost-utility analysis. There are two good reasons: they are not preference-based, and they do not incorporate death on the scale.

Quality-Adjusted Life Years

The concept of QALYs is displayed in Figure 1. The horizontal axis represents quantity of life (years), and the vertical axis represents quality of life (specifically, health-related quality of life). Consider a person who without the program would follow the bottom path and die at time death 1. With the program that person would follow the upper path and die later, death 2. The area between the two paths represents the health improvement achieved and is measured in QALYs.

One of the strengths of the QALY approach is that it can combine gains in quantity of life and gains in quality of life into a single metric. In Figure 1 the area marked A represents the gain in quality of life, and the area marked B is a gain in quantity of life. Specifically, area A is the gain in quality of life during the time the person would have been alive anyway, and area B is the gain in quantity of life (life extension) adjusted by the quality of that life extension.

Figure 1. Quality-Adjusted Life Years Gained from an Intervention



Adapted from Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine*. New York, NY: Oxford University Press, 1996, Figure 4.2.

Utilities

Theory. Modern utility theory was developed in 1944 by von Neumann and Morgenstern.[34] It is a theory of how individuals ought to make decisions in the face of uncertainty if they wish to act in a way that is defined as rational. The definition of rational behavior is provided by the fundamental axioms of utility theory. These axioms are simple statements that most people find compelling as statements of rational behavior in the face of uncertainty. The axioms and the system have remained the dominant normative paradigm for decision making under uncertainty for more than half a century.[35]

However, utility theory entails *ought* not *is*. It is not a behavioral or descriptive model of actual decision making under uncertainty in practice. An enormous amount of literature in psychology discusses actual decision-making behavior, and although the von Neumann-Morgenstern utility theory is indeed a good descriptive model in some cases,[36-38] generally it is not.[39,40] Nevertheless, this theory is still considered to be the best normative model.[35]

Direct Measurement of Utility. Detailed methods for direct measurement of utility are available elsewhere.[23,25,26] Some of the key points include the following. We generally recommend the use of marker states (eg, mild, moderate, severe) that relate to the types of health states under consideration. This provides some context for respondents in assessing their own health state. It also provides intermediate anchors for the scale, which is useful when rating other states. Although most textbooks de-

scribe utility measurement with examples in which death is the lower anchor, it is not necessary always to use death in this way. A living state can be used as the lower anchor for measurements; death is measured at least once but not more.

Our experience in using these instruments is that the visual analog scale is a useful warm-up exercise for respondents to become familiar with the states and to get them in the correct preferential order. It does not, however, provide cardinal utilities on an interval scale. For that task we prefer the standard gamble, which is founded directly on the axioms of utility theory. If the standard gamble cannot be used on all states, conversion functions are available to convert visual analog scale scores to standard gamble scores.[30]

Indirect Measurement of Utility. For most studies the indirect measurement of utilities with one of the established systems is easy and is quite sufficient. The systems require that the health status of the patient be recorded according to the classification system provided. A formula that comes with the system is used to calculate the preference score. The three best known systems at the moment are the quality of well-being,[27,28] the HUI (Health Utilities Index),[29-31] and the EuroQol-5D.[32,33] In all of these, the scoring formula is based on directly measured preferences of the general public. Thus the scoring formula is based on community preferences. This is seen as a strength of the systems because recent guidelines[11] recommend that the appropriate preferences for calculation of QALYs be community preferences.

History of Multi-Attribute Health Status Classification Systems. The quality of well-being was the first of these classification systems.[41] It contained four attributes, 4 to 35 levels per attribute, for a total of 3500 health states.

The HUI Mark 1 (HUI1) was described in 1982 and was based in part on the quality of well-being system.[42] The HUI1 contains four attributes, four to eight levels per attribute, and a total of 960 states. In 1984, this author and his colleagues undertook a study to identify the fundamental attributes that ought to be contained in such a system.[43] This provided six core attributes as follows: sensory and communication, physical activity, happiness, cognition, self-care, and pain and discomfort. The HUI2 is built directly on these attributes with the addition of one attribute—fertility—included because of the application we were studying at the time. The HUI2 consists of seven attributes, three to five levels per attribute, and a total of 24,000 health states.[44] Based on experience with the HUI2, we made some modifications and produced the HUI3.[29-31] It consists of eight attributes, five to six levels per attribute and a total of 972,000 health states.

In 1990 the first version of the EuroQol was described.[45] It contained six attributes, two to three levels per attribute and a total of 216 health states. The current version of EuroQol consists of five attributes, three levels per attribute, and a total of 243 health states.[32,33]

All these systems have a scoring formula based on directly measured preferences from the general public. However, the measurement instruments differ. The quality of well-being is based on visual analog scale measures, the EuroQol is based on time trade-off measures, and the HUI is based on standard gamble measures. Only the standard gamble measures have the advantage of providing von Neumann-Morgenstern utilities. The other instruments produce preferences but not utilities.

Health Utilities Index

Health Utilities Index 2 and Health Utilities Index 3. The two current versions of the HUI are the HUI2 and HUI3. Each version consists of a set of attributes and defined levels on each attribute ranging from no disability to major disability. Each version can be used descriptively as a health status classification system or numerically as a utility scoring formula. The two systems are slightly different from each other.

There are advantages and disadvantages of each, and our current recommendation for most users is to use both systems. It is particularly easy to use both systems, because there is a combined questionnaire that is simple to administer and captures the data to classify patients into both systems.

The classification systems for HUI2 and HUI3 are shown in Tables 1 and 2. The scoring formula for HUI2 is shown in Table 3. The questionnaire to be used in studies is available for various modes of administration (patient self-administration, interviewer administration face to face, interviewer administration by telephone), in two versions (current health and usual health), and in various languages. The appropriate questionnaire and procedures manuals are available from the Health Utilities Index group, McMaster University.

Table 1. Health Status Classification System for HUI2

Attribute	Level	Level Description
Sensation	1	Ability to see, hear, and speak normally for age
	2	Requires equipment to see or hear or speak
	3	Sees, hears, or speaks with limitations even with equipment
	4	Blind, deaf, or mute
Mobility	1	Able to walk, bend, lift, jump, and run normally for age
	2	Walks, bends, lifts, jumps, or runs with some limitations but does not require help
	3	Requires mechanical equipment (such as canes, crutches, braces, or wheelchair) to walk or get around independently
	4	Requires the help of another person to walk or get around and requires mechanical equipment as well
	5	Unable to control or use arms and legs
Emotion	1	Generally happy and free from worry
	2	Occasionally fretful, angry, irritable, anxious, depressed, or suffering night terrors
	3	Often fretful, angry, irritable, anxious, depressed, or suffering night terrors
	4	Almost always fretful, angry, irritable, anxious, depressed
	5	Extremely fretful, angry, irritable, or depressed usually requiring hospitalization or psychiatric institutional care
Cognition	1	Learns and remembers schoolwork normally for age
	2	Learns and remembers schoolwork more slowly than classmates as judged by parents and/or teachers
	3	Learns and remembers very slowly and usually requires special educational assistance
	4	Unable to learn and remember
Self-Care	1	Eats, bathes, dresses, and uses the toilet normally for age
	2	Eats, bathes, dresses, or uses the toilet independently with difficulty
	3	Requires mechanical equipment to eat, bathe, dress, or use the toilet independently
	4	Requires the help of another person to eat, bathe, dress, or use the toilet
Pain	1	Free of pain and discomfort
	2	Occasional pain; discomfort relieved by nonprescription drugs or self-control activity without disruption of normal activities
	3	Frequent pain; discomfort relieved by oral medicines with occasional disruption of normal activities
	4	Frequent pain, frequent disruption of normal activities; discomfort requires prescription narcotics for relief
	5	Severe pain; pain not relieved by drugs and constantly disrupts normal activities
Fertility*	1	Ability to have children with a fertile spouse
	2	Difficulty in having children with a fertile spouse
	3	Unable to have children with a fertile spouse

*Fertility attribute can be deleted if not required. Contact developers for details.
Source: Table II in Feeny D, Furlong W, Boyle M, Torrance G. Multi-attribute health status classification systems: Health Utilities Index. *Pharmacoeconomics* 1995;7:490-502. Used with permission from ADIS International.

Table 2. Health Status Classification System for HUI3

Attribute	Level	Level Description
Vision	1	Able to see well enough to read ordinary newsprint and recognize a friend on the other side of the street, without glasses or contact lenses
	2	Able to see well enough to read ordinary newsprint and recognize a friend on the other side of the street, but with glasses
	3	Able to read ordinary newsprint with or without glasses but unable to recognize a friend on the other side of the street, even with glasses
	4	Able to recognize a friend on the other side of the street with or without glasses but unable to read ordinary newsprint, even with glasses
	5	Unable to read ordinary newsprint and unable to recognize a friend on the other side of the street, even with glasses
	6	Unable to see at all
Hearing	1	Able to hear what is said in a group conversation with at least three other people, without a hearing aid
	2	Able to hear what is said in a conversation with one other person in a quiet room without a hearing aid, but requires a hearing aid to hear what is said in a group conversation with at least three other people
	3	Able to hear what is said in a conversation with one other person in a quiet room with a hearing aid, and able to hear what is said in a group conversation with at least three other people with a hearing aid
	4	Able to hear what is said in a conversation with one other person in a quiet room without a hearing aid, but unable to hear what is said in a group conversation with at least three other people even with a hearing aid
	5	Able to hear what is said in a conversation with one other person in a quiet room with a hearing aid, but unable to hear what is said in a group conversation with at least three other people even with a hearing aid
	6	Unable to hear at all
Speech	1	Able to be understood completely when speaking with strangers or friends
	2	Able to be understood partially when speaking with strangers but able to be understood completely when speaking with people who know the respondent well
	3	Able to be understood partially when speaking with strangers or people who know the respondent well
	4	Unable to be understood when speaking with strangers but able to be understood partially by people who know the respondent well
	5	Unable to be understood when speaking to other people (or unable to speak at all)
Ambulation	1	Able to walk around the neighborhood without difficulty, and without walking equipment
	2	Able to walk around the neighborhood with difficulty, but does not require walking equipment or the help of another person
	3	Able to walk around the neighborhood with walking equipment, but without the help of another person
	4	Able to walk only short distances with walking equipment, and requires a wheelchair to get around the neighborhood
	5	Unable to walk alone, even with walking equipment; able to walk short distances with the help of another person, and requires a wheelchair to get around the neighborhood
	6	Cannot walk at all
Dexterity	1	Full use of two hands and ten fingers
	2	Limitations in the use of hands or fingers, but does not require special tools or help of another person
	3	Limitations in the use of hands or fingers, is independent with use of special tools (does not require the help of another person)
	4	Limitations in the use of hands or fingers, requires the help of another person for some tasks (not independent even with use of special tools)
	5	Limitations in use of hands or fingers, requires the help of another person for most tasks (not independent even with use of special tools)
	6	Limitations in use of hands or fingers, requires the help of another person for all tasks (not independent even with use of special tools)
Emotion	1	Happy and interested in life
	2	Somewhat happy
	3	Somewhat unhappy
	4	Very unhappy
	5	So unhappy that life is not worthwhile

(Table 2 continued on next page)

Table 2. *(continued from previous page)*

Cognition	1	Able to remember most things, think clearly, and solve day-to-day problems
	2	Able to remember most things, but have a little difficulty when trying to think and solve day-to-day problems
	3	Somewhat forgetful, but able to think clearly and solve day-to-day problems
	4	Somewhat forgetful, and have a little difficulty when trying to think or solve day-to-day problems
	5	Very forgetful, and have great difficulty when trying to think or solve day-to-day problems
	6	Unable to remember anything at all, and unable to think or solve day-to-day problems
Pain	1	Free of pain and discomfort
	2	Mild to moderate pain that prevents no activities
	3	Moderate pain that prevents a few activities
	4	Moderate-to-severe pain that prevents some activities
	5	Severe pain that prevents most activities

Source: Table III in Feeny D, Furlong W, Boyle M, Torrance G. Multi-attribute health status classification systems: Health Utilities Index *Pharmacoeconomics* 1995;7:490-502. Used with permission from ADIS International.

Multi-Attribute Utility Theory. The Von Neumann-Morgenstern utility theory was extended by Keeney and Raiffa to cover consequences with multiple attributes.[46] For the extension, the authors had to add one assumption to the set of axioms underlying the von Neumann-Morgenstern utility theory. The additional assumption refers to the type of utility independence found, or assumed, among the attributes. Depending on the type of utility independence, the resulting multi-attribute utility function can be additive, multiplicative, or multilinear. Although the additive function is the simplest, our consistent finding is that the data do not support this formulation. The HUI2 scoring formula is based on a multiplicative function. In the HUI3 project both a multiplicative and a multilinear function are being investigated.

Health Utility Index Scoring. The utility scoring formula for HUI2 is shown in the caption under Table 3. The data for the formula is given in the table. To use the formula, the relevant *b*-score for each attribute is selected from the table based on the level of health status of the patient on that attribute, and the *b*-scores are plugged into the formula. The resulting score is a utility score for the patient's health state on a utility scale where dead has a utility of 0.00 and healthy has a utility of 1.00. A patient at level 1 on all attributes would score 1.00 overall. A patient at the worst level on all attributes would score -0.03 overall. Thus, such a

Table 3. Health Utilities Index Mark 2 Scoring Formula

Sensation		Mobility		Emotion		Cognition		Self-Care		Pain		Fertility	
x_1	b_1	x_2	b_2	x_3	b_3	x_4	b_4	x_5	b_5	x_6	b_6	x_7	b_7
1	1.00	1	1.00	1	1.00	1	1.00	1	1.00	1	1.00	1	1.00
2	0.95	2	0.97	2	0.93	2	0.95	2	0.97	2	0.97	2	0.97
3	0.86	3	0.84	3	0.81	3	0.88	3	0.91	3	0.85	3	0.88
4	0.61	4	0.73	4	0.70	4	0.65	4	0.80	4	0.64		
		5	0.58	5	0.53					5	0.38		

Formula: $u^* = 1.06(b_1 \times b_2 \times b_3 \times b_4 \times b_5 \times b_6 \times b_7) - 0.06$, where u* is the utility of the health state on a utility scale where dead has a utility of 0.00 and healthy has a utility of 1.00. Because the worst possible health state was judged by respondents as worse than death, it has a negative utility of -0.03. The standard error of u* is 0.015 for measurement error and sampling error, and 0.06 if model error is also included. x_i is attribute level code for attribute i; b_i is level score for attribute i.

patient is in a state judged to be worse than death.

The scoring formula and the data in Table 3 are based on a multiplicative multiattribute utility function with preferences measured among the general public with a standard-gamble instrument. We validated the scoring formula by comparing the results with directly measured utilities for health states that were not part of the development of the scoring formula. The HUI2 formula predicts these utilities with a standard deviation of prediction error of 0.06. Thus in using the HUI2 scoring formula one can add plus or minus 0.12 to the calculated score to determine a 95% confidence interval for a sensitivity analysis.

HUI3 scoring formulas are still under development. An HUI3 provisional scoring formula has been used in a number of studies, particularly studies of population health.[47-50] The HUI3 system also has a set of direct utility scores for 70 prevalent states.[51] Finally, the HUI3 multiplicative function has been completed but has not yet been fully tested or released, and the HUI3 multilinear function is still under development. As discussed later, it is important for users to appreciate that even without a scoring formula, these systems are extremely useful. We recommend that current users gather data for both the HUI2 and HUI3. For many studies that are starting now, the HUI3 formulas will be fully available by the time the data are to be analyzed.

Health Utility Index Applications. The HUI system is useful in two ways, descriptively and numerically. Descriptively, even without the scoring formula, the HUI system is an excellent classification system for recording and displaying the health status of cohorts of individuals, comparing among cohorts, and monitoring changes in cohorts over time. The HUI has been used in this way in both clinical studies[52-58] and population health studies.[47] Numerically, the HUI scoring formula gives a utility score that is appropriate as a composite measure of health-related quality of life. It is particularly well suited for use in calculating QALYs and quality-adjusted life expectancy. Again, this capability is useful in both clinical studies, particularly cost-utility studies,[59,60] and the measurement of population health.[48,50,61,62]

The HUI system is broadly applicable to a wide variety of patients and to the general public. In clinical studies it has been successfully used in an extremely diverse set of clinical conditions. In population health studies it has been used in a representative national sample of ambulatory and institutionalized members of the general public. In addition, it is being used in a growing number of countries. Translations are available or underway for a number of different languages.

Summary of Health Utility Index. The HUI is a comprehensive, generic, preference-based health status assessment and valuation system. It is based on community preferences, which are the recommended source of preferences for cost-utility analysis.[11] The scoring formula is founded directly on multi-attribute utility theory. The scores are provided on the conventional interval scale of health, in which death equals 0 and health equals 1. Population norm data are rapidly becoming available through the widespread use of the HUI in population health studies.

The HUI is useful in three kinds of studies. First, in clinical studies it can be used alongside the other clinical measures as a composite measure of health-related quality of life. Second, in economic evaluation studies it provides the preference-based score needed to calculate QALYs; that is, it enables the analyst to undertake a cost-utility analysis. Third, in population health applications it provides the score to calculate measures of population health such as quality-adjusted life expectancy, years of healthy life, health-adjusted person-years, and population health index.[47,50] The HUI system is easy to administer, taking only a few minutes of time to gather the necessary data. It is becoming

available in a variety of languages and is being widely adopted.

Cost-Utility Analysis

Research Agenda. The methods of cost-utility analysis itself is an active area of research. Researchers are investigating improvements in the methods of direct measurement of utilities. For example, interactive computer interviews and telephone administration are being investigated as replacements for the more expensive face-to-face interviewer administration. There is considerable work worldwide in furthering the development of multi-attribute systems such as the HUI. This includes investigations of more sophisticated scoring models such as the multilinear model. The work includes cultural and language adaptation for various countries and addition of disease-specific variations to the basic system for special areas such as musculoskeletal diseases, cardiovascular diseases, and mental health. Head-to-head comparisons of the various multi-attribute systems such as the HUI, quality of well-being, and EuroQol-5D are needed to determine the relative advantages and disadvantages of each. Theoretic issues such as whether conventional QALYs are reasonable approximations of utilities for paths of health states have to be further investigated. Indeed, the entire issue of whether path utilities are the standard needs further clarification. If path utilities are not the standard, then what is? A new concept based on person trade-offs is being advocated by some researchers as the more appropriate metric.[63,64] Basically, how do we know if we have done more good than harm at a population level? This is the fundamental question that requires more investigation. Captured in this question are issues of equity. For example, should QALYs be weighted differently for different individuals? In summary, there is no shortage of research to be done on cost-utility analysis, although the current methods are clearly suitable for use at present.

Conclusion

Cost-utility analysis is a special form of cost-effectiveness analysis. It was founded originally on decision analysis and operations research (optimization theory). Now it can be linked into a welfare economic foundation, or it can continue to be defined extra welfare. It uses the cost per QALY gained as the fundamental metric. The quality adjustment in the QALY should be preference-based, interval-scaled, and referenced to death. Appropriate methods for measuring such a quality adjustment include direct measurement with instruments such as the standard gamble or indirect measurement with systems such as the HUI.

··· REFERENCES ···

1. Torrance GW. *A Generalized Cost-Effectiveness Model for the Evaluation of Health Programs.* Buffalo, NY: State University of New York at Buffalo; 1971. Dissertation.
2. Torrance GW, Thomas W, Sackett D. A utility maximization model for evaluation of health care programs. *Health Serv Res* 1972;7:118-133.
3. Torrance GW. Health status index models: A unified mathematical view. *Manage Sci* 1976;22:990-1001.
4. Sinclair J, Torrance GW, Boyle M, et al. Evaluation of neonatal intensive care programs. *N Engl J Med* 1981;305:489-494.
5. Garber AM, Phelps CE. *Economic Foundations of Cost-Effectiveness Analysis.* Stanford, CA: National Bureau of Economic Research, 1995.
6. Garber AM, Weinstein MC, Torrance GW, Kamlet MS. Theoretical foundations of cost-effectiveness analysis. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine.* New York, NY: Oxford University Press, 1996:25-53.
7. Culyer A. The normative economics of health care finance and provision. *Oxf Rev Econ Policy* 1989;5:34-58.
8. Gold MR, Siegel JE, Russell LB, Weinstein MC. *Cost-Effectiveness in Health and Medicine.* New York, NY: Oxford University Press, 1996.
9. Torrance GW, Blaker D, Detsky A, et al. Canadian guidelines for economic evaluation of pharmaceuticals. *Pharmacoeconomics* 1996;9:535-559.

10. Menon D, Schubert F, Torrance GW. Canada's new guidelines for the economic evaluation of pharmaceuticals. *Med Care* 1996;34:DS77-DS88.
11. 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, NY: Oxford University Press, 1996:82-134.
12. Guyatt G, Veldhuyzen Van Zanten S, Feeny D, Patrick D. Measuring quality of life in clinical trials: A taxonomy and review. *Can Med Assoc J* 1989;140:1441-1448.
13. Feeny DH, Torrance GW, Labelle R. Integrating economic evaluations and quality of life assessments. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:85-95.
14. Clinch JJ. The functional living index—cancer: Ten years later. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:215-225.
15. Bellamy N, Buchanan W, Goldsmith CH, Campbell J, Stitt L. Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to anti-rheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833-1840.
16. Fretwell MD. Frail older patients: Creating standards of care. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:809-817.
17. Mangione CM, Phillips RS, Seddon JM, et al. Development of the Activities of Daily Vision Scale: A measure of visual functional status. *Med Care* 1992;30:1111-1126.
18. Mangione CM, Lee PP, Hays RD. Measurement of visual functioning and health-related quality of life in eye disease and cataract surgery. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. Philadelphia, PA: Lippincott-Raven, 1996:1045-1051.
19. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop* 1989;248:13-14.
20. Ware JE. The SF-36 health survey. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:337-345.
21. Damiano AM. The sickness impact profile. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:347-354.
22. McEwen J, McKenna SP. Nottingham health profile. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:281-286.
23. Torrance GW. Measurement of health-state utilities for economic appraisal: A review. *J Health Econ* 1986;5:1-30.
24. Bennett K, Torrance GW, Tugwell P. Methodologic challenges in the development of utility measures of health-related quality of life in rheumatoid arthritis. *Control Clin Trials* 1991;12:118s-128s.
25. Furlong W, Feeny D, Torrance GW, Barr R, Horsman J. *Guide to Design and Development of Health-State Utility Instrumentation*. Hamilton, Ontario. McMaster University, Centre for Health Economics and Policy Analysis, Working Paper No. 90-9, 1990.
26. Bennett KJ, Torrance GW. Measuring health state preferences and utilities: Rating scale, time trade-off and standard gamble techniques. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:253-265.
27. Kaplan R, Anderson J. A general health policy model: Update and applications. *Health Serv Res* 1988;23:203-235.
28. Kaplan RM, Anderson JP. The general health policy model: An integrated approach. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:309-322.
29. Feeny D, Furlong W, Boyle M, Torrance G. Multi-attribute health status classification systems: Health utilities index. *Pharmacoeconomics* 1995;7:490-502.
30. Torrance G, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions: Health utilities index. *Pharmacoeconomics* 1995;7:503-520.
31. Feeny DH, Torrance GW, Furlong WJ. Health utilities index. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:239-252.
32. Essink-bot M, Stouthard M, Bonsel G. Generalizability of valuations on health states collected with the EuroQol questionnaire. *Health Econ* 1993;2:237-246.
33. Kind P. The EuroQol instrument: An index of health-related quality of life. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1996:191-201.
34. von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1944.
35. Edwards W. Toward the demise of economic man and woman: Bottom lines from Santa Cruz. In: Edwards W, ed. *Utility Theories: Measurements and Applications*. Boston: Kluwer Academic, 1992:253-267.

36. Fischer GW. Utility models for multiple objective decisions: Do they accurately represent human preferences? *Decis Sci* 1979; 10:451-479
37. Currim I, Sarin R. Robustness of expected utility model in predicting individual choices. *Organ Behav Hum Decis Processes* 1992;52:544-568.
38. Nease RF. Do violations of the axioms of expected utility theory threaten decision analysis? *Med Decis Making* 1996;16:399-403.
39. Loomes G. Evidence of a new violation of the independence axiom. *J Risk Uncert* 1991;4:91-108.
40. Luce R. Where does subjective expected utility fail descriptively? *J Risk Uncert* 1992; 5:5-27.
41. Patrick D, Bush J, Chen M. Methods for measuring levels of well-being for a health status index. *Health Serv Res* 1973;8:228-245.
42. Torrance GW, Boyle MH, Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. *Oper Res* 1982;30:1043-1069.
43. Cadman D, Goldsmith C, Torrance GW, Boyle M, Furlong W. *Development of a Health Status Index for Ontario Children.* Final Report to Ontario Ministry of Health for Research Grant DM648 (00633). Hamilton, Ontario: McMaster University, Centre for Health Economics and Policy Analysis, 1986.
44. Feeny D, Furlong W, Torrance GW, Rosenbaum P, Weitzman S. A comprehensive multi-attribute system for classifying the health status of survivors of childhood cancer. *J Clin Oncol* 1992;10:923-928.
45. EuroQol Group. EuroQol: A new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199-208.
46. Keeney R, Raiffa H. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* New York, NY: Wiley, 1976.
47. Berthelot J, Roberge R, Wolfson M. The calculation of health-adjusted life expectancy for a Canadian province using a multi-attribute utility function: A first attempt. In: JM Robine, CD Mathers, MR Bone, I Romieu, eds. *Calculation of Health Expectancies: Harmonization, Consensus and Future Perspectives.* Colloque INSERM: 1992 Oct; Montpellier, France: John Libbey Eurotext, 1993:161-162.
48. Roberge R, Berthelot JM, Wolfson M. The health utility index: Measuring health differences in Ontario by socioeconomic status. *Health Rep* 1995;7:25-32.
49. Roberge R, Berthelot J, Wolfson M. *The Impact of Socioeconomic Status on Health Status in Ontario.* Ottowa, Ontario: Statistics Canada, 1993.
50. Wolfson MC. Health-adjusted life expectancy. *Health Rep* 1996;8:41-46.
51. Torrance GW, Feeny D, Furlong W, et al. Health utilities index mark 3: Utility scores for prevalent states. *Med Decis Making* 1995;15:435. Abstract.
52. Barr R, Furlong W, Dawson S, et al. An assessment of global health status in survivors of acute lymphoblastic leukemia in childhood. *Am J Pediatr Hematol Oncol* 1993;15:284-290.
53. Feeny DH, Leiper A, Barr R, et al. The comprehensive assessment of health status in survivors of childhood cancer: Application to high-risk acute lymphoblastic leukaemia. *Br J Cancer* 1993;67:1047-1052.
54. Barr R, Pai M, Weitzman S, et al. A multi-attribute approach to health status measurement and clinical management: Illustrated by an application to brain tumors in childhood. *Int J Oncol* 1994;4:639-648.
55. Barr R, Feeny D, Furlong W, Weitzman S, Torrance GW. A preference-based approach to health-related quality of life for children with cancer. *Int J Pediatr Hematol Oncol* 1995;2:305-315.
56. Barr R, Furlong W, Feeny D, et al. Evaluating treatments for childhood cancer. *Int J Technol Assess Health Care* 1995;11:1-10.
57. Gemke RJBJ, Bonsel GJ, van Vught AJ. Long term survival and state of health after paediatric intensive care. *Arch Dis Child* 1995;73:196-201.
58. Kiltie AE, Gattamaneni HR. Survival and quality of life of paediatric intracranial germ cell tumour patients treated at the Christie Hospital, 1972-1993. *Med Pediatr Oncol* 1995;25:450-456.
59. Boyle M, Torrance GW, Sinclair J, Horwood S. Economic evaluation of neonatal intensive care of very-low-birth-weight infants. *N Engl J Med* 1983;308:1330-1337.
60. Kelly AE, Haddix AC, Scanlon KS, Helmick CG, Mulinare J. Cost-effectiveness of strategies to prevent neural tube defects. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine.* New York, NY: Oxford University Press, 1996:313-348.
61. Tully P, Mohl C. Older residents of health care institutions. *Health Rep* 1995; 7:27-30.
62. Hood S, Beaudet M, Catlin G. A healthy outlook. *Health Rep* 1996;7:25-32.
63. Nord E. The person-trade-off approach to valuing health care programs. *Med Decis Making* 1995;15:201-208.
64. Nord E. Health status index models for use in resource allocation decisions: A critical review in the light of observed preferences for social choice. *Int J Technol Assess Health Care* 1996;12:31-44.