Special Article

Measurement of Clinical-Effect: Utility

Phantipa Sakthong PhD*

** Department of Clinical Pharmacy, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Bangkok*

The utility approach to assessing health-related quality of life is the most widely used technique for assessing preferences for health outcomes in the economic evaluation of health care. The scale for utility scores assigns a value of 1.0 to perfect health and 0.0 to death. The utility scores are employed to weigh time spent in each health state to estimate quality-adjusted life years (QALYs) gained, which is used as the denominator in cost-utility analysis and cost-effectiveness analysis. Utility scores are obtained through direct assessments using techniques such as standard gamble (SG), time-trade off (TTO), and visual analog scale (VAS), or by using multi-attribute systems such as the Health Utilities Index (HUI) or EuroQol (EQ-5D). According to international HE guidelines, the most preferred utility methods are SG and TTO, followed by EQ-5D, VAS and HUI, respectively. In Thailand, the EQ-5D is the most recommended utility method because it has acceptable feasibility and validity.

Keywords: *Utility, Preference, Health-related quality of life, Quality of life weights, Quality-adjusted life years, QALYs*

J Med Assoc Thai 2008; 91 (Suppl 2): S43-52
Full text. e-Journal: *http://www.medassocthai.org/journal*

Many diseases have a negative impact not only on patients' survival, but also on their health-related quality of life (HRQoL), which is not captured by conventional or biological clinical measures. HRQoL assessment has been extremely important for understanding the impact of diseases and treatments on patients' lives. It is a multidimensional construct including many health concepts, *e.g.* physical functioning, social and role functioning, mental health, and general health perceptions[1]. A common approach to measuring HRQoL is to employ generic and disease specific health status instruments[2,3]. The generic instruments can result in a single outcome score (health index) or a profile of scores (health profile). The index and the profile represent the two approaches to HRQoL assessment: the utility approach and the psychometric approach. Generic health profiles allow a determination of the effects of the treatment on different aspects of quality of life without necessarily using multiple instruments[4]. In addition, health profiles can be applied to a wide variety of conditions, so they allow comparisons of the effects on quality of life of different treatments in different diseases. However, generic health profiles have limitations. They may not be responsive to changes in specific conditions. For example, the items of the EuroQoL (EQ-5D) or SF-36 may not relate to symptoms that improve when antiretroviral drugs are used: fever, diarrhea, and weight loss. Another limitation of health profiles is that they do not produce a single preference score or value to calculate quality-adjusted life years (QALYs), so they cannot be used in cost-utility analysis (CUA). Utility measures of quality of life, such as standard gamble and EQ-5D, are another type of generic instrument. These measures are reported as a single index score. A major advantage of utility measurement is its application to CUA, which will be described in more details later.

The second approach to quality of life measurement focuses on aspects of health status that are specific to the area of primary interest such as disease-specific measures[4]. Specific measures therefore may be clinically sensible to the physician. The disadvantages of specific instruments are that they are not comprehensive and cannot be used to compare across conditions.

Correspondence to: Sakthong P, Faculty of Pharmaceutical Sciences, Chulalongkorn University, Prayathai Rd, Pathumwan, Bangkok 10330, Thailand. Phone: 0-2218-8408, Fax: 0-2218-8403, E-mail: Phantipa.S@chula.ac.th

Rationale of utility measurement and its theory

Utility is the value or worth placed on a level of health status, or improvement in health status, as measured by the preferences of individuals or society[5]. The utility measurement is necessary for the calculation of QALYs gained to ascertain the most commonly used health outcome measures in CUA and in cost-effectiveness analysis (CEA)[5,6]. Gold *et al* treat CUA as a specific type of CEA[6]. The QALY is a measure of life expectancy weighted by a utility score which is measured on a cardinal scale between 0 (death) and 1 (full health). It should be noted that the QALY also permits negative HRQoL values to represent health states worse than death[7]. In addition to application in CUA, a utility can be used for a clinical population to provide a single summary measure of HRQoL. This is because the utility score reflects both the health status and the value of that health status to the patient. Utilities can also be used as quality weights for calculating quality adjusted life expectancy as measures of population health[8]. There is no consensus regarding the most appropriate utility measurement approach. The recommendation of utility measurement in Thailand will be discussed in detail later.

The utility theory and its applications to health outcome measurement has its roots in the work of von Neumann and Morgenstern[9]. In 1944, John von Neumann, a mathematician, and Oscar Morgenstern, an economist, published their theory of rational decision-making under uncertainty, now called expected utility theory or von Neumann-Morgenstern utility theory. This decision theory described how a rational individual should make decisions when faced with uncertain outcomes. The utility approach that is based on this utility theory is called standard gamble (SG).

Sometimes the terms "utility", "value", and "preference" are interchangeably used; however, they differ. Preference is the overall concept and has two different types: utility and value. There are two important aspects of the measurement process. One is the way in which the question is framed, and if the outcomes are certain or uncertain. The other is the way in which a subject is asked to respond: scaling or making a choice. The methods of measuring utilities are shown in Table 1[5]. The first dimension of the measurement process is question framing. A question framed under certainty asks the subject to compare two or more outcomes and to choose between them or to scale them. The outcome is certain and has no probabilities. A question framed under uncertainty asks the subject to compare two alternatives where at least one of the alternatives is uncertain. This outcome contains probabilities. The difference between the two methods of questioning is that whereas the certainty method does not capture the subject's risk attitude, the uncertainty method does. In the real world future health outcomes are uncertain, so the utility method is a more appropriate measurement method than the value method.

The second aspect of the measurement process is the response method. A subject can be asked to determine their strength of preference by giving a number on a numerical scale. Alternatively, a subject can be asked to choose between two alternatives. The first approach is rooted in psychology or psychometric scaling, while the second method comes from economics and decision sciences. Many analysts prefer the choice-based method.

In summary, the methods in cells 1 and 3 measure values, while those in cell 4 measure utilities (Table 1). The difference between cells 3 and 4 is risk attitude. The difference between cells 1 and 3 is the difference between choosing and scaling. The details of each method are described in the next topic.

Utility Methods

Directly measured utility methods

There are three well-known methods of directly measuring utilities namely Visual Analog Scale (VAS), SG, and Time trade-off (TTO).

Visual analogue scale

The VAS is a common rating scale approach which is based on the information integration theory, which explains the cognitive process of judgment[10]. This theory includes two constructs: integration and valuation. Other rating scales include the rating scale (RS) and the category scale (CS). The RS refers to a

Table 1. Methods of measuring utilities

Response method	Question framing	
	Certainty (values)	Uncertainty (utilities)
Scaling	1 Rating scale Category scaling Visual analogue scale	2
Choice	3 Time trade-off	4 Standard gamble

scale of numbers, often 0-100. The CS contains a number of categories, often 0-10. The VAS shows a respondent a line, often 10 cm in length, with defined endpoints such as "death" at the lower bound and "perfect health" at the upper bound. The respondents are asked to mark the point on the scale to indicate their value of health. The line can vary in length, and be vertical (mostly) or horizontal. The VAS approach is the least difficult direct method of assessing utility, and takes less respondent time.

Standard gamble

The SG method is based on the utility theory of decision making under uncertainty proposed by von Neumann and Morgenstern[9], and is the original method of measuring utility[11]. Using the SG, utility is determined by the choices respondents make as the probabilities of immediate death or full health are varied. For example, respondents are offered two alternatives between 1) living in a health state with HIV/AIDS for the rest of their life and 2) taking a risky treatment. If the treatment were successful, the patients would live in perfect health. However, if the treatment failed, they would die immediately, with no pain. The probability of treatment success (p) is varied until the respondent is indifferent towards the two alternatives. The utility of the respondents is the probability (p) they chose. The SG approach is graphically presented in Fig. 1.

Time trade-off

The TTO is theoretically appealing because it is conceptually equivalent to a QALY. It was developed as an alternative to SG and was designed to overcome the problem of explaining probabilities to respondents[12]. In a TTO, subjects are asked about the number of years in a health state (t) they would be willing to tradeoff for a shorter life span in full health (x). The utility is x/t. For example, the respondents are given a choice between living with HIV for 10 years or living for a lesser number of years in perfect health. The utility is calculated by dividing the lesser number of years by 10 years.

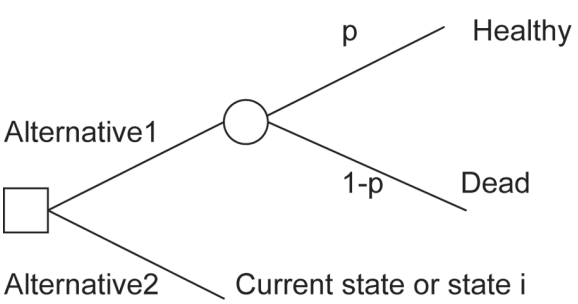

Fig. 1 Standard Gamble for a health state preferred to death

Indirectly measured utility methods

An alternative method commonly used indirectly derives utility via a multi-attribute health status classification system. This method involves a two-step procedure and makes use of health status instruments such as quality of well-being (QWB), EQ-5D, and the health utilities index (HUI). Respondents are asked to rate their level of a particular health state with several attributes e.g. mobility and pain. These attribute levels are then mapped to a 0-to-1 quality of life scale using weights that were previously elicited from the community by the instrument developers. Using these methods, respondents indirectly assign their health states to a quality of life weight or utility. Hence, these are called indirect methods of utility measurement.

Quality of well-being

The QWB scale includes 4 attributes: mobility, physical activity, social activity, and the symptom-problem complex[13]. There are 3 categories on the mobility and physical activity scales and 5 on the social activity scale. The symptom-problem complex has 27 symptoms. The scoring function is based on CS measurements from a random sample of the general public. The resulting scoring function is between 0.00-1.00. The QWB scale is time consuming to use and code. Its estimated time to complete is 15-18 minutes, but a shorter version is available.

EuroQoL

The EQ-5D includes 5 attributes: mobility, self care, usual activity, pain/discomfort, and anxiety/depression[14]. Each attribute has 3 levels: no problem, some problems, and major problem. The scoring function was measured using the TTO method on a random sample of approximately 3,000 adults in the United Kingdom[15,16]. The resulting score was between -0.59-1.00. The estimated time of completion was about 1 minute. The EQ-5D has been translated into many languages including Thai (see Appendix).

Health utilities index

The most common HUI methods are HUI2 and HUI3[17,18]. The HUI2 method was initially applied to childhood cancer. Subsequently, the HUI2 method has

been modified for adult applications. The HUI2 consists of 7 attributes: sensation, mobility, emotion, cognition, self-care, pain, and fertility. Each attribute has 4-5 levels, varying from highly impaired to normal. In addition, the fertility attribute can be easily dropped from both the classification system and the scoring formula if not needed. The HUI2 scoring function was measured on a random sample of parents of schoolchildren in Hamilton, Canada using both VAS and SG methods. The resulting score was between -0.03-1.00.

The HUI3 method was based closely on the HUI2 method. The fertility attribute, however, was dropped, and the sensory attribute was expanded into three attributes: vision, hearing, and speech. The HUI3 consists of 8 attributes: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain with 5 or 6 levels per attribute. The HUI3 scoring function was measured on a random sample of general adult population in Hamilton, Canada using both VAS and SG methods. The final version of HUI3 employed the SG method[19]. The resulting score was between -0.36-1.00. The questionnaire takes about 10 minutes for self-administration and 2-3 minutes for interviewer administration.

SF-6D

The SF-36 was reconstructed into a six dimensional measure called the SF-6D[20]. A sample of 249 SF-6D health states were valued by 611 subjects, considered to be representative of the UK population, using the SG method. Models were developed to predict health state evaluations for all possible health states defined by the SF-6D. Brazier et al devised an algorithm for estimating utilities from existing SF-36 data via their SF-6D classification. The SF-6D was able to explain about 58% of the variance in the SG scores. The resulting score was between 0.46 - 1.00.

Other methods for deriving utilities

In addition to directly and indirectly measuring utilities methods, other approaches that can be used to derive utilities are as follows:

Expert opinion

Expert opinion can be employed to estimate utilities. It should, however, only be used when no other data sources exist or when the parameters are of secondary importance in the analysis. In addition, it should be elicited in a structured manner, such as the Delphi method[21].

Mapping VAS to SG or TTO

Since VAS is easier to administer, cheaper and less time consuming than SG and TTO, there is an attempt to map VAS to SG or TTO. However, there is a lack of evidence of a stable relationship, so SG or TTO scores should be obtained directly rather than estimating the scores from VAS.

Willingness to pay

Willingness to pay (WTP) is a technique for assessing preference for health outcomes of public health and healthcare interventions. WTP estimates are used to assess the value of heath gain in monetary values. In the WTP method, patients are asked how much they are willing to pay to be free of the disease, where a higher WTP indicates a worse quality of life. The method used to measure WTP is known as the contingent valuation method[22].

Comparisons of the international economic evaluation guidelines

Twenty-nine guidelines from 24 countries are included in the International Society of Pharmacoeconomics and Outcome Research (ISPOR) health economic (HE) guidelines including 22 HE guidelines, 6 submission guidelines for formulary listing and one for journal publication[23]. The countries that provide HE guidelines are as follows: Austria, Australia, the Baltic region (Latvia, Lithuania, Estonia), Belgium, Canada, China, France, Germany, Hungary, Ireland, Israel, Italy, Netherlands, New Zealand, Norway, Poland, Portugal, the Russian Federation, Scotland, Spain, Sweden, Switzerland, England & Wales, and the United States of America (USA).

The utility approaches that many countries (9 countries) prefer to use for economic evaluation are SG and TTO methods. These countries include Poland, Belgium, France, Switzerland, Sweden, England & Wales, Italy, Portugal and China. EQ-5D ranks as the third preferred utility method. There are countries which employ the EQ-5D: Hungary, Poland, Sweden, Portugal, New Zealand and the Baltic Region. There are 4 countries that allow the use of the VAS or the rating scale. These countries are Poland, France, Switzerland, and China. Poland and the Baltic region are the only 2 countries that recommend the use of HUI. In addition, France permits the use of a WTP approach, but it must be justified. None of the countries recommend using expert's opinions. Most of the countries do not give the reasons for selecting their preferred utility methods, but a number of countries have described that the SG

and the TTO have been chosen because they are choice-based methods or based on the utility theory, which is the sacrifice of length of life for better health, while the VAS is not appropriate because it measures value rather than utility.

Nevertheless, many countries (10 countries) do not state specifically which utility methods should be applied for HE evaluations. These countries are: the Russian Federation, Germany, the Netherlands, Austria, Ireland, Scotland, Spain, Canada, Norway, and the U.S.A. The guidelines of Canada and Austria state that the selection should be justified and those of Scotland and Ireland state that the utility methods need to be outlined and qualified. Spain and the USA have just said that generic measures should be used without specifying the preferred methods. Similarly, the Netherlands has said that direct and indirect methods should be utilized but it did not specify which methods. Germany has said that the index and profile approaches should be employed. The Russian Federation has not stated anything about the preferred utility methods.

Specifically, Norway has proposed transforming utility to value (not the same as value, which is another type of preference) and supplementing this with cost-value analysis. Nord, a well-known Norwegian health economist, has reported that valuing health programs in terms of QALYs disregards societal concerns for fairness in resource allocation[24]. Thus, obtaining estimates of value incorporating concerns of fairness, based on the degree of severity of the illness, allows a more comprehensive and valid cost-value analysis of health care.

The submission guidelines of Belgium, Israel, and the USA do not state the preferred method to derive utility, while those of the British Medical Journal, Australia, and England & Wales say that the details of the methods used need to be given. Only Canada's submission guidelines originating from the Canadian Coordinating Office for Health Technology Assessment (CCOHTA) state a preference for the use of indirect approaches including HUI, EQ-5D, and QWB. Most guidelines suggest contemporary use of valid and reliable generic and disease specific instruments for assessing HRQoL.

Recommendations for Thai Health Technology Assessment (HTA) Guidelines

Based on international HE guidelines, the most preferred utility methods are SG and TTO, followed by EQ-5D, VAS or RS, HUI, and QWB, respectively. Which utility method should be recommended in Thailand? The following criteria for determining performance of a utility method can be used[25].

1. Practicality: Its acceptability to respondents. Such acceptability can be a function of length of time, complexity and respondents' interest in the task.

2. Reliability: The ability of a measure to reproduce the same values on separate administrations when there has been no change in the health state being valued.

3. Validity: The extent to which an instrument measures what it is intended to measure.

4. Responsiveness: Its ability to measure changes in health.

According to the psychometric criteria above, for Thai people I would recommend EQ-5D as the most preferred utility method. The reasons for choosing the EQ-5D are as follows:

In terms of practicality, compared to other direct methods (SG, TTO) and indirect methods (HUI, QWB, SF-6D), EQ-5D is shorter and easier to administer and to understand. The SG and TTO techniques are quite difficult and need well-trained interviewers. As for the HUI, QWB, and SF-6D, they are also time-consuming to administer, and thus causing respondent burden. As for VAS, it is viewed as the least difficult direct method. My study (n = 120) compared the performance of EQ-5D, VAS, and SG. It was found that EQ-5D was slightly easier than VAS but the difficulty rating scores were not significantly different ($p > 0.05$)[26]. One reason why the EQ-5D method was easier than the VAS method was that it provided interviewees with the dimensions of health, so they did not need to integrate them like using the VAS method.

Regarding validity, SG is undoubtedly the most theoretically appealing of the utility techniques because it has foundations in the expected utility theory which is the dominant theory of decision-making under uncertainty[9]. However, there is evidence that respondents violate the axioms of the utility theory (such as risk attitude)[27]. There is also evidence showing that SG values can be influenced by the frame of the gamble and the manner in which the task is presented[28]. The specific probabilities that are used may influence the SG scores as well[29].

An alternative method to the SG method is the TTO method. There is evidence to suggest that duration effects (a period of time of a health state) and time preference effects (the rate at which a decision maker is willing to trade a present for a future outcome) can have an impact on the elicitation of TTO values[30]. Elderly people under severe conditions often refuse to

trade off length of life because they place a greater value on survival than on their quality of life[31,32]. Some reasons for refusing to trade life time in exchange for health improvements are expressed in the following words: "I choose to live day by day", "time with my family is too precious", "the question is too hypothetical so older patients failed to grasp what was asked." In addition, both SG and TTO may be influenced by indifferent point searching procedures[33].

Even though VAS is not a choice-based technique like the SG and the TTO methods, it does not consider attitude toward risk or incorporate time horizons. The rating task can be also influenced by upper and lower bounds[34]. Subjects also tend to shy away from using the ends of the scale[35]. A rating task depends on a subject's numeric or quantitative reasoning skills[36]. If respondents have little experience with rating their health in relation to numbers, they may perform rating tasks like VAS poorly. In addition, cognitive abilities and emotions can also be a threat to the validity of the elicitation of SG, TTO, and VAS[37].

Because of the problems described above of the direct methods of deriving utilities, there is an alternative approach: discrete-state health index models or the multi-attribute health status classification system. As described before, this system works by attaching fixed utility weights to observable health states. Respondents are not required to make judgments or decisions about their utilities. Commonly used methods include EQ-5D, HUI, and QWB. Most international guidelines prefer EQ-5D to HUI or QWB. Also, the most frequently used instrument for calculating QALYs based on actual measurements of patients' HRQoL is EQ-5D[38].

Moreover, the EQ-5D method has acceptable reliability, validity, and responsiveness[39-41], even though some studies reported that the EQ-5D method had poor responsiveness[42]. It also has the problem of high ceiling effects[43,44]. The high ceiling effects and lack of sensitivity to change problems may be caused by having only three degrees of severity in the EQ-5D's dimensions. I have been told that the EuroQoL group is developing a new version of EQ-5D with five levels of severity. For now, one approach to coping with these problems is to develop a disease-specific utility instrument. This method can estimate the EQ-5D scores from disease-specific measures or map clinical data to the utility values. However, the disease-specific utility instrument does not allow for comparison across diseases, only among different strategies within the same disease.

A summary of the recommendations of utility methods in Thailand is as follows:

1. If a researcher collects a primary data of utility, EQ-5D is the most recommended utility method (the Thai algorithm version is preferred when available). Other algorithms that we can use include the UK[15], US[45], and Japanese[46] versions. Other direct and indirect utility methods such as VAS, SG, TTO, HUI, QWB, SF-6D, WTP can also be used but should be justified. The utility method selected should have data supporting the practicality, reliability, validity, and responsiveness among Thai people.

2. For a secondary data of utility, if there are Thai utilities available, the Thai utilities should be used. But if Thai data are not available, utilities from other population groups, whose characteristics are similar to those of the Thai people, should be applied. Alternatively, a systemic approach including meta-analysis should be employed to combine utilities taken from different studies. It is also recommended that they subject the results to sensitivity analyses to utilities.

3. Expert opinion, mapping VAS to TTO and SG are not recommended.

4. Disease-specific measures should be used contemporarily with utility measures.

5. A perspective of utility measurement depends on the objectives of the study. Use a patient's perspective when making treatment decisions for the individual patient; use a societal perspective (general public) when making program funding or policy decisions and generating Thai population-based utility weights for indirectly measured utility methods. A proxy can also be used when the subjects such as children, the elderly, and disabled cannot answer. A proxy can be parents, family members, or care givers. Whichever perspective or subject is applied, we must be sure that the subjects are well-informed and unbiased.

Acknowledgement

The funding support for this guideline development was provided through the Health Intervention and Technology Assessment Program (HITAP) by the Thai Health Promotion Foundation, the Health System Research Institute and the Bureau of Health Policy and Strategy, Ministry of Public Health.

References

1. Bungay KM, Boyer JG, Steinwald AB, Ware JE. Health-related quality of life: an overview. In: Bootman JL, Townsend RJ, McGhan WF, editors. Principle of pharmacoeconomics. Cincinnati, OH:

Wharvey Whitney Books; 1996: 128-48.
2. Cramer J, Spilker B. Quality of life and pharmacoeconomics: an introduction. Philadelphia: Lippincott-Raven; 1998.
3. Coons SJ, Kaplan RM. Cost-utility analysis. In: Bootman JL, Townsend RJ, McGhan WF, editors. Principle of pharmacoeconomics. Cincinnati, OH: Wharvey Whitney Books; 1996: 102-26.
4. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. Ann Intern Med 1993; 118: 622-9.
5. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. New York: Oxford University Press; 2005.
6. Gold MR, Russell LB, Siegel JE, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.
7. Patrick DL, Starks HE, Cain KC, Uhlmann RF, Pearlman RA. Measuring preferences for health states worse than death. Med Decis Making 1994; 14: 9-18.
8. Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions. Health utilities index. Pharmacoeconomics 1995; 7: 503-20.
9. von Neumann J, Morgenstern O. Theory of games and economic behavior. Princeton, NJ: Princeton University Press; 1944.
10. Anderson NH. Contributions to information integration theory. Hillsdale, NJ: Erlbaum Publishers; 1991.
11. Torrance GW, Feeny D. Utilities and quality-adjusted life years. Int J Technol Assess Health Care 1989; 5: 559-75.
12. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. Health Serv Res 1972; 7: 118-33.
13. Kaplan RM, Anderson JP. A general health policy model: update and applications. Health Serv Res 1988; 23: 203-35.
14. Brooks R. EuroQol: the current state of play. Health Policy 1996; 37: 53-72.
15. Dolan P, Gudex C, Kind P, Williams A. A social tariff for EuroQol: results from a UK general population survey, Discussion Paper No. 138. York: Centre for Health Economics, University of York; 1995.
16. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Health Econ 1996; 5: 141-54.
17. Feeny D, Furlong W, Boyle M, Torrance GW. Multi-attribute health status classification systems: Health Utilities Index. Pharmacoeconomics 1995; 7: 490-502.
18. Torrance GW, Furlong W, Feeny D, Boyle M. Multi-attribute preference functions: health utilities index. Pharmacoeconomics 1995; 7: 503-20.
19. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002; 40: 113-28.
20. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002; 21: 271-92.
21. Gustafson DH, Fryback DG, Rose JH, Yick V, Prokop CT, Detmer DE, et al. A decision theoretic methodology for severity index development. Med Decis Making 1986; 6: 27-35.
22. Diener A, O'Brien B, Gafni A. Contingent valuation studies in the health care literature: a review and classification. Health Econ 1998; 7: 313-26.
23. Tarn TYH, Smith MD. Pharmacoeconomic guidelines around the world. 2004 [cited 2007 March 24]. Available from: http://www.ispor.org/PEguidelines/index.asp.
24. Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. Health Econ 1999; 8: 25-39.
25. Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. Health Econ 1999; 8: 41-51.
26. Sakthong P. Health-state utilities and health-related quality of life in patients with HIV/AIDS in Thailand [dissertation]. Minneapolis: University of Minnesota Twin Cities; 2006.
27. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. J Health Econ 1996; 15: 209-31.
28. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. The measurement of patients' values in medicine. Med Decis Making 1982; 2: 449-62.
29. Wakker P, Stiggelbout A. Explaining distortions in utility elicitation through the rank dependent model for risky choices. Med Decis Making 1995; 15: 180-6.
30. Dolan P, Gudex C. Time preference, duration and health state valuations. Health Econ 1995; 4: 289-99.
31. Tsevat J, Dawson NV, Wu AW, Lynn J, Soukup JR, Cook EF, et al. Health values of hospitalized

patients 80 years or older: HELP Investigators: Hospitalized Elderly Longitudinal Project. JAMA 1998; 279: 371-5.
32. Arnesen TM, Norheim OF. Quantifying quality of life for economic analysis: time out for time trade off. J Med Ethics 2003; 29: 81-6.
33. Lenert LA, Cher DJ, Goldstein MK, Bergen MR, Garber A. The effect of search procedures on utility elicitations. Med Decis Making 1998; 18: 76-83.
34. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. Med Care 2000; 38: 583-637.
35. Bleichrodt H, Johannesson M. An experiment test of the theoretical foundation for a rating scale valuations. Med Decis Making 1997; 17: 208-16.
36. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med 1997; 127: 966-72.
37. Lenert L, Kaplan RM. Validity and interpretation of preference-based measures of health-related quality of life. Med Care 2000; 38(9 Suppl): II138-50.
38. Rasanen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryynanen OP, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: a systematic literature review. Int J Technol Assess Health Care 2006; 22: 235-41.
39. Krabbe PF, Peerenboom L, Langenhoff BS, Ruers TJ. Responsiveness of the generic summary measure compared to the disease-specific EORTC QLQ C-30. Qual Life Res 2004; 13: 1247-53.
40. Bushnell DM, Martin ML, Ricci JF, Bracco A. Performance of the EQ-5D in patients with irritable bowel syndrome. Value Health 2006; 9: 90-7.
41. Schweikert B, Hahmann H, Leidl R. Validation of the EuroQol questionnaire in cardiac rehabilitation. Heart 2006; 92: 62-7.
42. Wu AW, Jacobson KL, Frick KD, Clark R, Revicki DA, Freedberg KA, et al. Validity and responsiveness of the euroqol as a measure of health-related quality of life in people enrolled in an AIDS clinical trial. Qual Life Res 2002; 11: 273-82.
43. Luo N, Chew LH, Fong KY, Koh DR, Ng SC, Yoon KH, et al. Validity and reliability of the EQ-5D self-report questionnaire in English-speaking Asian patients with rheumatic diseases in Singapore. Qual Life Res 2003; 12: 87-92.
44. Marra CA, Woolcott JC, Kopec JA, Shojania K, Offer R, Brazier JE, et al. A comparison of generic, indirect utility measures (the HUI2, HUI3, SF-6D, and the EQ-5D and disease-specific instruments (the RAQoL and the HAQ) in rheumatoid arthritis. Soc Sci Med 2005; 60: 1571-82.
45. Shaw JW, Johnson JA, Coons SJ. US valuaion of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005; 43: 203-20.
46. Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ 2002; 11: 341-53.

Appendix

แบบสอบถามคุณภาพชีวิต EuroQoL (Thai version)

โปรดกาเครื่องหมาย √ ลงในกล่อง ☐ ที่แสดงถึงภาวะทางสุขภาพของข้าพเจ้าในวันนี้ได้มากที่สุด

1. *ความสามารถในการเคลื่อนไหว*
   - ☐ *ข้าพเจ้าไม่มีปัญหาเกี่ยวกับการเดิน*
   - ☐ *ข้าพเจ้ามีปัญหาเกี่ยวกับการเดินบ้าง*
   - ☐ *ข้าพเจ้าไม่สามารถเดินได้ จำเป็นต้องนอนอยู่บนเตียง*

2. *การดูแลตนเอง*
   - ☐ *ข้าพเจ้าไม่มีปัญหาในการดูแลร่างกายด้วยตนเอง*
   - ☐ *ข้าพเจ้ามีปัญหาบ้างในการใส่เสื้อผ้าหรืออาบน้ำด้วยตนเอง*
   - ☐ *ข้าพเจ้าไม่สามารถใส่เสื้อผ้าหรืออาบน้ำด้วยตนเอง*

3. *การทำกิจวัตรประจำวัน (เช่น การทำงานหาเลี้ยงชีพ, การเรียน, การทำงานบ้าน, การทำกิจกรรมกับครอบครัว, หรือการทำงานอดิเรก)*
   - ☐ *สุขภาพของข้าพเจ้าไม่มีผลต่อการทำกิจวัตรประจำวันดังกล่าวข้างต้น*
   - ☐ *สุขภาพของข้าพเจ้ามีผลบ้างต่อการทำกิจวัตรประจำวันดังกล่าวข้างต้น*
   - ☐ *สุขภาพของข้าพเจ้ามีผลทำให้ข้าพเจ้าไม่สามารถทำกิจวัตรประจำวันดังกล่าวข้างต้น*

4. *ความเจ็บปวด/ความไม่สบาย*
   - ☐ *ข้าพเจ้าไม่มีอาการปวดหรือรู้สึกไม่สบาย*
   - ☐ *ข้าพเจ้ามีอาการปวดหรือรู้สึกไม่สบายปานกลาง*
   - ☐ *ข้าพเจ้ามีอาการปวดหรือรู้สึกไม่สบายอย่างมาก*

5. *ความวิตกกังวล/ความซึมเศร้า*
   - ☐ *ข้าพเจ้าไม่มีความวิตกกังวลหรือความซึมเศร้า*
   - ☐ *ข้าพเจ้ามีความวิตกกังวลหรือความซึมเศร้าปานกลาง*
   - ☐ *ข้าพเจ้ามีความวิตกกังวลหรือความซึมเศร้าอย่างมาก*

การวัดอรรถประโยชน์

พรรณทิพา ศักดิ์ทอง

วิธีการวัดอรรถประโยชน์ (utility) เพื่อใช้ประเมินคุณภาพชีวิตที่เกี่ยวกับสุขภาพเป็นวิธีการที่ใช้กันอย่างแพร่หลายมากที่สุดเพื่อหาความพึงพอใจ (preferences) ของผลลัพธ์ทางด้านสุขภาพในการประเมินความคุ้มค่าทางการแพทย์ของการดูแลทางด้านสุขภาพ คะแนนอรรถประโยชน์จะมีค่าอยู่ระหว่าง 0-1 โดยที่ 1 เท่ากับภาวะสุขภาพที่สมบูรณ์และ 0 เท่ากับการเสียชีวิต คะแนนอรรถประโยชน์จะนำมาใช้ในการคูณกับช่วงเวลาที่ใช้ในแต่ละสภาวะทางสุขภาพเพื่อใช้สำหรับการคำนวณหาปีสุขภาวะที่เพิ่มขึ้น (quality-adjusted life years gained) ซึ่งเป็นตัวหารในการวิเคราะห์ต้นทุนอรรถประโยชน์และการวิเคราะห์ต้นทุนประสิทธิผล คะแนนอรรถประโยชน์สามารถหาได้โดยวิธีการประเมินทางตรงเช่น วิธี standard gamble (SG), time-trade off (TTO) และ visual analog scale (VAS) และโดยวิธีทางอ้อม เช่นการใช้แบบสอบถาม Health Utilities Index (HUI) และ EuroQol (EQ-5D) ตามคู่มือการประเมินเทคโนโลยีด้านสุขภาพของประเทศต่างๆ พบว่าวิธีการวัดอรรถประโยชน์ที่ได้รับความนิยมมากที่สุดคือวิธี SG และ TTO ตามมาด้วยวิธี EQ-5D, VAS, HUI ตามลำดับ สำหรับประเทศไทยวิธีของการใช้แบบสอบถาม EQ-5D น่าจะเป็นวิธีที่เหมาะสมมากที่สุดเนื่องจากวิธีการนี้มีความเป็นไปได้และมีความเที่ยงตรงที่สามารถยอมรับได้