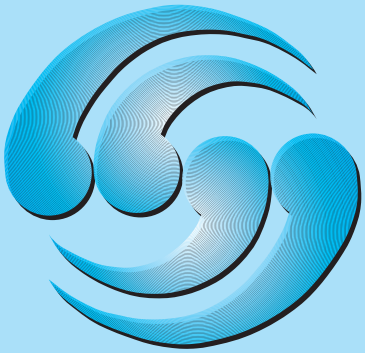
Volume 1, number 4



What are health utilities?

Sponsored by an educational grant from Aventis Pharma

Stavros Petrou
BSc MPhil PhD
Health Economist, National Perinatal Epidemiology Unit, University of Oxford

- **Utilities are cardinal values** that represent the strength of an individual's preferences for specific outcomes under conditions of uncertainty. They provide an approach to the measurement of health-related quality of life.

- Measuring health utilities involves three steps: **defining a set of health states** of interest, **identifying individuals** to provide judgements of the desirability of each health state and **aggregating across the individuals** to determine scale values for each health state.

- The issues that need to be addressed in any utility study include ensuring that the **relevant attributes and levels of health** are incorporated into the measurement instrument, determining the **appropriate preference-scaling** method to elicit utilities, deciding the **appropriate subjects** whose preferences should be evaluated and **minimising the impact of context effects** on the results.

- The methods that have been used to collect data on utilities include the **standard gamble approach**, the **time trade-off approach**, **rating scales** and the **willingness-to-pay** approach.

- Utilities have been used as the **preference weights** (quality levels) within the quality-adjusted life year (QALY) model.

www.evidence-based-medicine.co.uk

Abbreviated prescribing information is on page 6

What are utilities?

The term utilities (as currently used by health economists) refers to ***cardinal values that represent the strength of an individual's preferences for specific outcomes under conditions of uncertainty***.

Specifically, health utilities are ***preferences for specific health states or treatments. They provide an approach to the comprehensive measurement of health-related quality of life***.

The general approach to measuring health utilities includes three steps:

- Defining a set of health states of interest.
- Identifying individuals to provide judgements of the desirability of each health state.
- Aggregating across the individuals to determine scale values for each health state.[1]

Within this general framework, however, there are a number of issues that need to be addressed. These issues include, but are not limited to:

1. What are the relevant health dimensions?
2. What preference-scaling method should be used?
3. Whose preferences should be measured?
4. How may the impact of context effects be minimised?

The standard gamble approach in practice

An individual might be asked to choose between the certainty of surviving for a fixed period in a particular state of ill health and a gamble between surviving for the same period without disability on the one hand and immediate death on the other. The probability of surviving without disability, as opposed to dying, is then varied until the person shows no preference between the certain option and the gamble. This probability then defines the utility of an individual for the disabled state on a scale between 0 and 1, whose end-points are death and perfect health.

1. What are the relevant health dimensions?

The single most crucial component of any study that measures health-related quality of life is ensuring that the **relevant dimensions and levels of health are incorporated** into the measurement instrument. If these dimensions and levels do not correspond to patient outcomes, it will not be possible to detect differences even if such differences exist.

Within the context of utility measurement, **health dimensions are referred to as attributes**. Examples of health attributes include: physical, social and cognitive function, psychological wellbeing, symptoms and pain. In general, the maximum number of attributes that may be included in a health state is nine (preferably fewer), based on research that indicates humans can simultaneously process only 5–9 pieces of information.[2]

2. What preference-scaling method should be used?

There are a number of scaling methods that have been used to collect data on utilities including **standard gamble, time trade-off, rating scale, equivalence technique, ratio scaling** and **willingness to pay**.

Standard gamble

The standard gamble approach is the classic method of measuring preferences in economics, first presented by von Neumann and Morgenstern.[3] It uses hypothetical lotteries as a means of measuring people's preferences. These lotteries involve a choice between two alternatives: the certainty of one outcome and a gamble with two possible outcomes (see Box, left).

Breyer and Fuchs[4] used the standard gamble approach to investigate the risk behaviour of adults in the health dimension. A sample of 325 individuals was questioned about attitudes to 12 hypothetical scenarios.

In each scenario the individuals were offered choices between two alternative forms of treatment for a supposed illness. One treatment offered a fixed health effect. The other offered two possible outcomes (a more favourable and a less favourable health effect).

The authors found that individuals exhibit risk-averse behaviour towards positive health outcomes (they tend to opt for the certain and intermediate outcome) and risk-seeking behaviour towards negative health outcomes (they are willing to take more of a risk in order to have a chance of experiencing the most positive outcome on offer).

Petrou and Campbell[5] used the standard gamble approach to estimate utilities for a range of health states in colorectal carcinoma. They were able to demonstrate that the quality-of-life benefits of stabilisation in the treatment of advanced metastatic colorectal cancer were rated almost as highly as those of partial response. The study also showed that the benefits of irinotecan (Campto®), a drug licensed for the treatment of metastatic colorectal cancer in patients who had failed an established 5-FU-containing regimen, outweighed the short-term impact of toxicity in those patients who achieved at least stabilisation of their disease.

Advantages and disadvantages of common preference-scaling methods

While each scaling method has its advocates, none has been conclusively proven superior. The advantage of the rating scale approach is its efficiency, straightforwardness, wide applicability and the fact that it is quick and inexpensive to employ. However, there is a temptation for subjects to spread their responses evenly across categories.

The standard gamble remains the gold standard for many health economists because it is based on decision-making under conditions of uncertainty, a component of most studies in healthcare. However, this approach is relatively time-consuming and people often have difficulties understanding the concept of probabilities. Moreover, utilities or preference values can be strongly influenced by the way questions are framed.[6]

Despite the fact that the time trade-off approach has advantages over standard gamble and is reliable, practicable, and a reliable measure of utility, it does have some drawbacks. As Rosser and Kind[7] point out, the approach is based on the assumptions that the perception of time is linear and that the perception of the severity of illness is independent of the time spent in this state. In addition, the trade-off concept is difficult for many people to understand.

Time trade-off

The time trade-off approach was developed as a simpler alternative to standard gamble. It involves asking subjects to consider the relative amounts of time they would be willing to trade in order to survive in various health states. The choice may lie between continuing in a present defined state of ill health or moving to a shorter but healthier life. The duration of survival in the healthier state is varied until the subject indicates no preference between the two alternatives, at which point his/her utility for the health state can be calculated.

Rating scale

The rating scale is based on psychometric theory. It consists of a single line with anchors representing best possible health and death (or some alternative). Respondents are asked to place each health state on the line such that the intervals between the placements reflect their perceived differences between the health states. Rating scale estimates are consistently lower than those obtained via the standard gamble or time trade-off methods. Rating scales are often used to introduce the standard gamble exercise because they allow the respondent to become familiar with the health states and with ranking.

Two additional techniques, equivalence technique and ratio scaling, have also occasionally been used to measure health-state utilities, but they are not widely used and will not be discussed here.

Quality levels in measurement

All these approaches yield utilities that range from 0–1, where a higher score represents better health-related quality of life. These utilities have been widely used as the preference weights (quality levels) within the quality-adjusted life year model. Quality-adjusted life years (QALYs) are single-weighted measures that combine the survival periods

There is evidence that utilities differ among different population groups

and health-related quality-of-life states accruing from health programmes. They provide a standard unit for measuring health gain across diseases and specialties (see Box, below).

Often the QALY health gain derived from interventions is combined with the cost of delivering those interventions in a cost–utility framework. Because district health authorities will already be providing many of the services, this is usually expressed as the marginal cost per QALY – the marginal cost of increasing provision of an intervention to a level that secures one extra QALY.

Willingness-to-pay

Willingness-to-pay provides an alternative approach to measuring health utilities. Rather than measuring utilities on a 0–1 scale, this approach estimates utility in monetary terms. The benefits of a healthcare programme are estimated on the basis of the total amount individuals are willing to pay for the health improvement itself. Likewise, compensation for a diminution in health status can also be estimated by use of, for example, the calculation of pay premiums received by workers in dangerous occupations.

One advantage of this approach is that people are familiar with answering willingness-to-pay questions for goods or benefits in everyday life. Willingness-to-pay estimates can be obtained in a number of ways depending on the context. Examples include questioning consumers themselves through the use of questionnaires, directly observing their behaviour in the marketplace, inferring from their behaviour and the use of models.

However, there are a number of practical difficulties involved in measuring willingness-to-pay. Importantly, the approach assumes that people employ rational considerations when responding to willingness-to-pay questions and produce responses that are a function of ability to pay.

3. Whose preferences should be measured?

This is a question that must be addressed in all studies of health-related quality of life. Utilities are usually obtained from clinicians or other experts, patients and the general public.

- The **primary justification for using clinicians or other healthcare professionals** is that they are knowledgeable about the health states and may be more accessible than some patient groups – thus they may serve as reasonable proxies for patients.
- The **primary reason for selecting patients** is that they are the people who experience the impact of the disease and treatment, so it is their preferences that should be considered of most importance. Thus studies of new therapies are most often conducted from this perspective.
- The **rationale for measuring the preferences of the general public** is related to policy decision-making – that is, decisions pertinent to public policy should be based on public opinion.[8]

The question of whose preference to measure is important because there is some evidence that utilities differ among different population groups. Sackett and Torrance[9] reported that the health state of the respondent was related to utilities for some health states, but Llewellyn-Thomas *et al*[10] found that the rater's own health status did not influence ratings and Carter *et al*[11] found that health professionals tended to assign lower ratings than the general public.

Evidence from the geriatric literature strongly suggests that there are differences between patient and proxy respondents.[12] In general, the correlation between patient and proxy responses varied by health dimension, but proxy ratings tended to be lower than patient ratings.

Quality-adjusted life years in practice

Let us assume that there are two treatments for an illness. Both treatments extend the life expectancy of an individual by eight years. However, treatment A results in the individual surviving the years in full health (usually represented by a utility score of 1 on a cardinal scale), while treatment B results in the individual surviving the years in a health state with a utility score of 0.5. Treatment A has led to a gain in QALYs of 8 (8 x 1) while treatment B has led to a gain in QALYs of 4 (8 x 0.5).

4. How may the impact of context effects be minimised?

Most inconsistencies in measurement of health utilities are the result of differences in perspective that may arise from the way the health state or treatment is presented to the respondent. Such effects may result from:

- **Use of specific terms referred to as 'labelling effects'** – for example, referring to a common treatment for cancer as 'radiation' decreased its probability of selection from when it was referred to as a 'treatment with specific outcomes'.[9]
- Research suggests that **anchor effects may occur where ratings change** depending, for example, on whether the anchor of death or worst possible health is used.[13]
- The **form in which the health state is presented** to the respondent has also been shown to affect ratings.
- **Prognosis and duration of health state** have been shown to affect preference ratings.[14]

However, almost nothing is known about the impact of culture on health utilities.

These inconsistencies are called context effects and may be minimised by careful attention to construction and presentation of the health states.

The following guidelines have been suggested for development and administration of health states to avoid context effects:[14–16]

- Use health states of the same duration.
- Avoid inclusion of diagnosis, disease labels, laboratory test results and prognosis in health-state descriptions.
- Use trained personnel to collect data.
- Pose questions to the respondent in an unbiased manner to reduce framing effects.
- Use a variety of administration techniques to reduce cognitive burden (for example, props that allow presentation of the task in a clear and concise manner and presentation of subsets where the total number of health states is large).

Note

For further information on utility or willingness-to-pay techniques, references 3–6 are particularly recommended.

References

1. Torrance GW, Boyle MH, Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. *Oper Res* 1982; **310:** 1043–1069.
2. Miller GA. The magical number seven plus or minus two; some limits on our capacity to process information. *Psychol Rev* 1956; **63:** 81–97.
3. von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior.* New York: John Wiley, 1953.
4. Breyer F, Fuchs VR. *Risk Attitudes in Health Care: An exploratory study*. Cambridge, MA: NBER Working Paper No 875, 1982.
5. Petrou S, Campbell N. Stabilisation in colorectal cancer. *Int J Palliat Nurs* 1997; **3**(5): 275–280.
6. Read JL, Quinn RJ, Berwick DM, Fineberg HV, Weinstein MC. Preferences for health outcomes: comparisons of assessment methods. *Med Decis Making* 1984; **4**(3): 315–329.
7. Rosser RM, Kind P. A scale of valuations of states of illness: is there a social consensus? *Int J Epidemiol* 1978; **7**(4): 347–358.
8. Drummond M, Stoddart G, Torrance G. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford Medical Publications, 1987.
9. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. *J Chronic Dis* 1978; **7:** 347–358.
10. Llewellyn-Thomas H, Sutherland HJ, Trichler DL *et al*. Benign and malignant breast diseases: the relationship between women's health status and health values. *Med Decis Making* 1991; **11:** 180–188.
11. Carter WB, Bobbitt RA, Bergner M *et al*. Validation of an interval scaling: the Sickness Impact Profile. *Health Serv Res* 1976; **11**(4): 516–528.
12. Rothman ML, Hedrick SC, Bulcroft KA *et al*. The validity of proxy-generated scores as measures of patient health status. *Med Care* 1991; **29**(2): 115–124.
13. Kaplan RM, Ernst JA. Do rating scales produce biased preference weights for a health index? *Med Care* 1983; **21:** 193–207.
14. Froberg DG, Kane RL. Methodology for measuring health-state preferences – III: population and context effects. *J Clin Epidemiol* 1989; **42:** 585–592.
15. Froberg DG, Kane RL. Methodology for measuring health-state preferences – I: measurement strategies. *J Clin Epidemiol* 1989; **42:** 345–354.
16. Froberg DG, Kane RL. Methodology for measuring health-state preferences – II: scaling methods. *J Clin Epidemiol* 1989; **42:** 459–471.

Abbreviated prescribing information: Campto® (irinotecan hydrochloride trihydrate)

Presentations: Vials of concentrate for infusion containing either 40 mg or 100 mg irinotecan hydrochloride trihydrate. **Indications:** Treatment of adult patients with advanced colorectal cancer. **Dosage & Administration:** Solution must be prepared aseptically. Campto should be administered as an intravenous infusion over 30 to 90 minutes. In first line: combination therapy of 180 mg/m^2 every 2 weeks followed by folinic acid and 5-fluorouracil; in second line: monotherapy 350 mg/m^2 every 3 weeks. Prophylactic anti-emetics are recommended. ***Dosage Adjustments:*** Subsequent cycles should follow appropriate recovery of all adverse events to grade 0 or 1 NCI-CTC and resolution of diarrhoea. Dose reduction of 15-20% recommended, if patients experience grade 4 neutropenia, febrile neutropenia, thrombocytopenia, grade 4 leucopenia or grade 3-4 non-haematological toxicity. ***Impaired hepatic function:*** Monitor liver function frequently. Do not use where bilirubin > 1.5 x ULN. ***Impaired renal function:*** Not recommended. ***Elderly:*** Care due to the greater frequency of decreased biological functions. **Contraindications:** Chronic inflammatory bowel disease and/or bowel obstruction; severe hypersensitivity reactions to Campto; pregnancy; breastfeeding; severe bone marrow failure; WHO performance status > 2. **Warnings and Precautions:** Use in units specialised in the administration of cytotoxic chemotherapy and under the supervision of an oncologist. Patients needing closer follow-up or particular risk of neutropenia weekly dosing schedule (125 mg/m^2/week for 4 weeks, then 2 weeks rest) may be considered. Patients should be aware of the risks of acute cholinergic syndrome and neutropenia, and management of delayed diarrhoea (occuring > 24 hours after the infusion). Loperamide should not be given prophylactically. Weekly monitoring of full blood counts recommended. Patients should not drive if dizziness or visual disturbances occur. Contraceptive measures must be taken during and for 3 months after therapy. **Interactions:** Care in patients receiving neuromuscular blocking agents. **Adverse reactions:** Delayed diarrhoea (requires immediate treatment with loperamide). Uncommonly, pseudomembranous colitis. Neutropenia, fever, anaemia, thrombocytopenia, nausea and vomiting, acute cholinergic syndrome. Infrequently intestinal obstruction, ileus or gastrointestinal haemorrhage, and intestinal perforation. Transient increases in transaminases, alkaline phosphatase, bilirubin or creatinine. Dyspnoea, muscular contractions, cramps, paraesthesia, asthenia, reversible alopecia, dehydration, constipation; infrequently dehydration-related renal insufficiency, hypotension or circulatory failure. Mild effects include anorexia, cutaneous reactions, abdominal pain, and mucositis. Uncommonly, allergy and infusion site reactions. **Pharmaceutical Precautions:** Do not mix with any other medications. Complete infusion within 12 hours of reconstitution, if stored at room temperature (22 ± 4°C) or 24 hours, if stored at 2–8°C. Comply with prevailing cytotoxic handling guidelines when preparing or handling Campto.
Legal category: POM PL Number: 40 mg: 0012/0302, 100 mg: 0012/0303. **Basic NHS Price:** Campto 40 mg; £53.00; Campto 100 mg; £130.00.
Further information is available on request from Aventis Pharma Ltd, 50 Kings Hill Avenue, West Malling, Kent. ME19 4AH. **Date of Preparation:** October 2000.

This publication, along with the others in the series, is available on the internet at www.evidence-based-medicine.co.uk

Sponsored by an educational grant from *Aventis Pharma*

Any enquiries please contact:
Team Assistant to the Health Economics Unit
Aventis House
Kings Hill
West Malling
Kent
ME19 4AH
Tel: 01732 584 254
Fax: 01732 584 029

Published by Hayward Medical Communications, a division of Hayward Group plc.

ONC 137117
Date of preparation: May 2001