

Utilities and QALYs in health economic evaluations: glossary and introduction

Gianni Virgili · Daniela Koleva · Livio Garattini · Rita Banzi · Gian Franco Gensini

Received: 25 May 2010 / Accepted: 16 June 2010 / Published online: 6 July 2010
© SIMI 2010

Systematic reviews of evidence: from effectiveness to efficiency

Economics studies the optimal allocation of resources for the production of benefits to the society. By definition, resources are limited, and healthcare systems must face this condition every day. This may explain why economic evaluations are increasingly used as a key component of decision making. For instance, they represent a mainstay in any Health Technology Assessment (HTA) exercise and are subject to a guidelines' development process [1]. After the assessment of the efficacy and safety of health interventions, economic evaluations enter to compare alternative options in terms of costs and consequences. Full

economic evaluations consist of a comparison of both costs (resource use) and consequences (outcomes, effects) of alternative strategies [2].

If Cochrane reviews are intended to help providers, practitioners and patients in making informed decisions about health care interventions, they cannot afford to ignore their economic aspects. Cochrane reviews assemble, select, critique and combine reliable data from multiple research studies on the effectiveness and other issues related to healthcare interventions. The latest version of the Cochrane Handbook for Systematic Reviews of Interventions [3], freely accessible at <http://www.cochrane-handbook.org>, acknowledges that incorporating economic perspectives and evidence into Cochrane reviews can enhance their usefulness and applicability for healthcare decision-making. An entire chapter (“Incorporating economics evidence”), provides guidance on how authors of systematic reviews can collect the best evidence on the economic aspects of interventions in addition to the best evidence on their clinical effectiveness [4].

A leading point here is to expand our knowledge on efficiency (optimal use of resources) after having established effectiveness (whether treatments work) [5].

The main types of full economic evaluations (cost-effectiveness, cost-utility and cost-benefit analyses) compare alternative interventions in terms of both costs (resource use) and consequences (effects) and can be included in Cochrane reviews, as well as studies that only focus on costs and resource use, without comparing health alternatives, i.e. partial economic evaluations [4].

Cost-effectiveness analyses measure the effects of an intervention and its comparators in identical units of outcome (e.g. mortality), thus comparing them in terms of ‘cost per unit of effect’ [5]. In a cost-benefit analysis both resource use and effects of alternative interventions are

G. Virgili (✉)
Department of Specialised Surgical Sciences, University of Florence and Azienda Ospedaliero-Universitaria Careggi, Florence, Italy
e-mail: gianni.virgili@unifi.it

D. Koleva · L. Garattini
CESAV, Center for Health Economics, “Mario Negri” Institute for Pharmacological Research, Ranica BG, Italy

R. Banzi
Italian Cochrane Centre, Mario Negri Institute for Pharmacological Research, Milan, Italy

G. F. Gensini
Department of Critical Care Medicine and Surgery, University of Florence and Azienda Ospedaliero-Universitaria Careggi, Florence, Italy

G. F. Gensini
Don Carlo Gnocchi Foundation, Onlus IRCCS, Impruneta FI, Italy

expressed in monetary units. Cost-utility analysis weighs the effect of interventions by means of people's preferences on the trade-off between length of life and subjective levels of well-being associated with health states. The best known utility measure is the quality-adjusted life year (QALY). Alternative interventions are compared in terms of cost per unit of utility gained (Cost per QALY).

Cost-utility analyses alongside randomized controlled trials (RCTs) or as a part of modelling studies comprise a common type of economic evaluation. The methodology of measurement of effects, utilities and QALYs in such analyses is the focus of this paper.

Utilities and QALYs in health economic evaluations

The incremental cost-effectiveness ratio in the case of cost-utility analyses includes in its denominator the concept of QALYs which, in turn, is tightly related to the term "utility". Utility is usually a number between 0 and 1 ("0" representing the worst health state and "1"—perfect health). The conventional use of these utility values is to convert them into QALYs, by their multiplication by the years spent in certain health states. The aim of QALYs is to combine the survival and health-related quality of life (QoL) into a single measure, allowing thus the comparison of cost-utility estimates of different interventions for the purposes of public health decision making.

How utilities are obtained

Several methods to measure the utilities exist. Probably the most successful one implies a questioning of a nationally representative sample of people to trade life duration with life quality [6].

One of the techniques, the Time Trade-Off (TTO) approach, consists in asking people how much of their life expectancy they would exchange for perfect health, if they were allowed to live without a specific disease from now onwards. For example, if persons suffering from chronic pain are ready to give back 4 out of 10 years they expect to live, they actually score the disease utility as 0.6 (1–0.4). The utility values assigned to various health states depend on who collects the information through predesigned questionnaires—health professionals or lay people.

The second approach called Standard Gamble (SG), asks patients to choose among two health states: usually a chronic condition lived in a limited health and the alternative to receive a treatment that leads to instant death or to complete recovery. Accepting, for example, a 15%

probability of immediate death would mean the assignment of a utility value of 0.85 (1–0.15) to that health state.

Another approach, based on patient's rating promptness, is a visual-analogue scale (VAS), in which participants are asked to express the magnitude of their preferences by selecting a point on a line (typically graduated from 0 to 100) with definite anchors (endpoints)—death and perfect health. The VAS has not been accepted by some HTA agencies since it does not carry the implication of asking people to trade life duration with QoL [7].

When a specific physical function is assessed, the endpoints may differ from the scheme death-perfect health. In the case of visual function assessment, endpoints may include blindness and perfect vision, trying to choose between utility loss associated with eye disease, and that associated with other causes. In a study, the adjustment of utility scores for age and comorbidity ensures that utility loss differences attributable to different grades of eye disease are caused by visual disease and not by other factors [8]. However, the authors state that such an adjustment does not allow one the ability to distinguish the utility loss caused by visual disease in patients with mild severity, nor to distinguish between visual and nonvisual dysfunction in a multiple organ disease like diabetes. Probably the importance of the endpoint perfect vision in the assessment of ophthalmic disease utilities is due to the emphasis on comorbidity as a confounding factor, and the study shows that changing the endpoints leads to further problems [8].

Multi-attribute utility instruments (MAUIs)

Utilities can also be associated with health states defined by means of generic health-related QoL questionnaires [7]. In the case of MAUIs, the utility obtained through TTO or SG methods is linked to a QoL instrument based on statistical modelling of a scoring algorithm. The most common MAUIs are EQ-5D, SF-6D (allowing a transformation of SF-36 or SF-12 scores into QALY weights) and the Health Utility Index (HUI 2 and HUI3) [7]. They are generic instruments for measurement of health-related quality of life aimed at economic evaluations; for example, the EQ-5D comprises five domains related to the following areas: mobility, self-care, routine daily activities, pain/discomfort, and anxiety/depression.

The advantage of this method is that the questionnaires can be used during RCTs, allowing one to estimate QALYs for both the experimental intervention and control group, and then to combine these estimates with costs obtained alongside the trials.

Measuring QALYs gained comparing an intervention with a control group

Once utilities (QALY weights) are estimated, QALYs are further calculated by multiplying the person's stay in a health state by its utility value. The cost per QALY gained thanks to the intervention over a certain time period is the difference between the cost per QALY ratio of the intervention and that of the control group; the information on efficacy, safety and costs of both intervention and control groups is derived from RCTs or even from multiple sources.

How QALYs are used in economic evaluations

Among the various measures employed in economic evaluations, the Cost-Effectiveness Acceptability Curve (CEAC) plots the probability of a treatment to be cost-effective (on the y axis) as a function of increasing levels of willingness-to-pay (Fig. 1) [8]. CEACs play a crucial role in economic evaluations, e.g. those developed for the UK National Institute for Clinical Excellence (NICE). Due to the numerous assumptions unavoidable in economic modelling, the robustness of the economic evaluation results should be tested through sensitivity analysis. The assumptions are usually related to

1. utilities or cost estimates;
2. efficacy over time: relative efficacy of two treatment options is generally estimated for the first few years after a treatment initiation in RCTs, but economists usually project treatment benefits to the limits of a patient's average life expectancy;

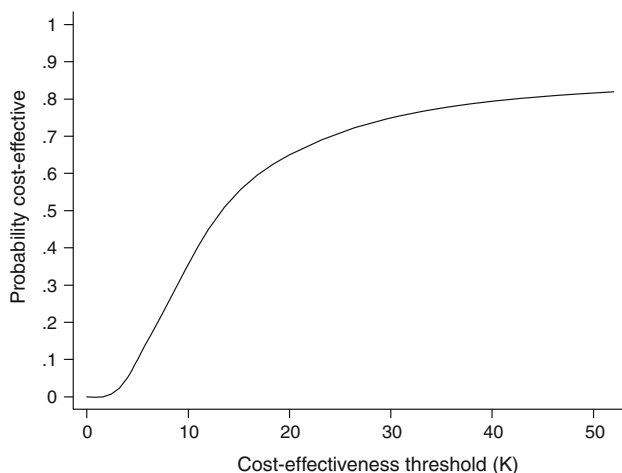


Fig. 1 The Cost-Effectiveness Acceptability Curve (CEAC) plots the probability of a treatment to be cost-effective (on the y axis) as a function of increasing levels of willingness-to-pay (on the x axis)

3. discount rate: generally costs and benefits are discounted by 3 to 5% for a time horizon above 2 years.

The use of flexible Bayesian statistical economic models can overcome the need to perform several one-way sensitivity analyses since they incorporate data uncertainty, starting from the results of a meta-analysis of RCTs, and proceeding with the intrinsic uncertainty of economic modelling [9].

Limitations in the use of utilities in economic evaluations: transferability of results

A recent review by Knies et al. [12] discusses the international transferability of utilities derived from EQ-5D questionnaires. The authors find substantial differences between national EQ-5D value sets, and discourage the uncritical application of utilities from other countries to the individual setting. These large differences are due not only to national cultural behaviours throughout the set of countries examined, but also reflected the methodological differences between them, which made the speculations in the field very difficult.

The International Society for Pharmaco-economics and Outcomes Research (ISPOR) has reviewed different national guidelines with regard to their statements about transferability, and finds that less than 40% of guidelines include recommendations about the transferability of health-state evaluations or utility estimates across countries [2]. It is generally believed that costs always differ between countries and therefore, should always be adjusted. Differences in clinical effectiveness parameters between countries can also lead to differences in economic evaluations' results [10].

Should we rely on economic evaluations?

A reviewer trying to examine economic aspects of interventions through a systematic review, or a reader interested in the critical appraisal of an economic evaluation, should consider the following aspects: If the economic question is clearly stated: i.e. what is the economic burden of a health condition to the society? What are the incremental resource consequences of implementing an intervention versus a comparator? What is the economic value associated with a benefit, or adverse events as a result from the intervention? Does the economic question stated determine the methodology, time horizon and perspective of the analysis similar to what happens in any other research field? Useful data to collect on the characteristics of each economic study are year of the study; details of interventions and comparators; study design and source(s) of resource use, unit costs and

(if applicable) effectiveness data; decision-making jurisdiction, geographical and organizational setting; viewpoint of the analysis; and time horizon for which the costs and effects are assessed.

As for results of the economic evaluation, estimates of specific cost items associated with both intervention and the comparator and the related unit costs estimates (if reported) should be extracted separately, and subsequently the resource quantities should be multiplied by the respective unit costs. The type and quantity of each resource used should be registered as “natural units” (e.g. length of hospital stay in days, duration of operation in minutes, number of outpatient attendances at 6-month follow-up, number of days of work). It is also important to collect information on the price year and currency used to estimate the costs. Measures of incremental resource use and costs should possibly be collected at an individual patient level. Moreover, sensitivity analysis and any information regarding the impact of varying assumptions on the magnitude and direction of final results should be conducted.

The issue of methodological quality of economic evaluations has been raised frequently [11–14]. The critical appraisal of health economics studies aims to assess whether they describe aspects like methods, model assumptions, and possible bias sources in a transparent and fully supported manner by the available evidence, in order to facilitate any critical reader [13].

A number of checklists have been developed to critically appraise the quality of economic evaluations. However, no checklists have been formally validated, and in many cases they do not fully extend over all possible bias domains. Two checklists have become important: the British Medical Journal Checklist for authors and peer reviewers of economic submissions [15], and the CHEC list for assessment of methodological quality of economic evaluations [16]. Guidelines for conducting economic evaluations are available from several institutions, and efforts are made to find an agreement on methods to be used [2]. The lack of clarity of the economic evaluations in reporting information on crucial items affects the judgements on their methodological quality (<http://www.equator-network.org/>).

In conclusion, it is worth quoting an invited commentary entitled “The futility of utilityweighting” [17], which acknowledged the artificial judgment imposed by the complex reality, and quoted Churchill in reference with the democracy, concluding that it seems we have an imperfect system but, nonetheless, the best we can create at the moment.

Conflict of interest None.

References

- Guyatt GH, Oxman AD, Kunz R, Jaeschke R, Helfand M, Liberati A, Vist GE, Schünemann HJ (2008) GRADE Working Group. Incorporating considerations of resources use into grading recommendations. *BMJ* 336(7654):1170–1173
- Drummond M, Barbieri M, Cook J et al (2009) Transferability of economic evaluations across jurisdictions: ISPOR good research practices task force report. *Value Health* 12:409–418
- Higgins JPT, Green S (eds) (2008) *Cochrane handbook for systematic reviews of interventions*. Wiley, Chichester
- Shemilt I, Mugford M, Byford S, Drummond M, Eisenstein E, Knapp M, Mallender J, McDavid D, Vale L, Walker D (2008) Incorporating economics evidence, Chap 15. In: Higgins JPT, Green S (eds) *Cochrane handbook for systematic reviews of interventions*, Version 5.0.1 (updated September 2008). The Cochrane Collaboration, 2008. Available from <http://www.cochrane-handbook.org>
- Cochrane AL (1971) *Effectiveness and efficiency: random reflections of health services*, 2nd edn. Nuffield Provincial Hospitals Trust, London (published 1989)
- Drummond MF, O’Brien B, Stoddart GL, Torrance GW (1997) *Methods for the economic evaluation of health care programmes*, 2nd edn. Oxford University Press, Oxford
- Scuffham PA, Whitty JA, Mitchell A, Viney R (2008) The Use of QALY weights for QALY calculations. A review of industry submissions requesting listing on the Australian pharmaceutical benefits scheme 2002–2004. *Pharmacoeconomics* 26(4):297–310
- Lee BS, Kymes SM, Nease RF, Sumner W, Siegfried CJ, Gordon MO (2008) The impact of anchor point on utilities for 5 common ophthalmic diseases. *Ophthalmology* 115:898–903
- Rawlins MD, Culyer AJ (2004) National Institute for Clinical Excellence and its value judgments. *BMJ* 329:224–227
- Canadian Agency for Drugs and Technologies in Health (2006) *Guidelines for the economic evaluation of health technologies: Canada*, 3rd edn. Canadian Agency for Drugs and Technologies in Health, Ottawa
- Drummond MF, Sculpher MJ, Torrance GW et al (2005) *Methods for the economic evaluation of health care programmes*, 3rd edn. Oxford University Press, Oxford
- Knies S, Evers SMAA, Candel MJJM, Severens JL, Ament AJHA (2009) Utilities of the EQ-5D. Transferable or not? *Pharmacoeconomics* 27(9):767–779
- Neumann PJ, Greenberg D, Olchanski NV, Stone PW, Rosen AB (2005) Growth and quality of the cost-utility literature, 1976–2001. *Value Health* 8:3–9
- Rennie D, Luft HS (2000) Pharmacoeconomic analyses: making them transparent, making them credible. *JAMA* 283:2158–2160
- Drummond MF, Jefferson TO (1996) Guidelines for authors, peer reviewers of economic submissions to the BMJ The BMJ Economic Evaluation Working Party. *BMJ* 313:275–283
- Evers S, Goossens M, de Vet H, van Tulder M, Ament A (2005) Criteria list for assessment of methodological quality of economic evaluations: consensus on Health Economic Criteria. *Int J Tech Assess Health Care* 21:240–245
- Kane RL (2008) The futility of utility weighting. *J Clin Epidemiol* 61:1195–1196