

## Reference Case<sup>1</sup>

A reference case is a set of methodological practices intended to enable, by means of standardization, meaningful comparisons of economic evaluation results both within and across different diseases and interventions. Such comparisons are unavoidable if economic analyses are expected to inform health care resource allocation decisions.

A reference case may be interpreted as a specific, highly prescriptive variant of a methodological guideline for health economic evaluations. Methodological guidelines have been developed as tools to support the conduct of scientifically consistent economic studies. Informal guidelines developed by academic groups often are differentiated from formalized guidelines issued by official bodies charged with technology appraisals to inform reimbursement and pricing decisions.

## Background

In the absence of a standard, analysts were free to make choices, including (but not limited to) the form of evaluation method (e.g., cost benefit versus cost effectiveness analysis), the appropriate measure of benefit (e.g., willingness to pay versus health outcomes), the perspectives for valuation (i.e., the source of preference data, e.g., patients versus a representative sample of the general public, individual versus social, ex ante versus ex post, or the choice of scaling instrument for utility measurement, such as standard-gamble, time-trade-off, person-trade-off, etc.) and costing (e.g., from a payer's

---

<sup>1</sup>Quote as: **Michael Schlander: Reference Case.** In: Michael Kattan (Editor) & Mark E. Cowen (Associate Editor), **Encyclopedia of Medical Decision Making.** Thousand Oaks, CA: SAGE 2009 (in press)

or from a societal viewpoint), the discounting of future benefits and costs, and the reporting of their findings. The resulting variation of analytic approaches would greatly decrease the policy value of economic analyses. In response, the concept of reference case analysis was proposed to serve as a point of comparison by a common core of methodological choices across studies. It is widely acknowledged that reference case analysis, although prescriptive and generic (i.e., not disease-specific) by definition, should not prevent analysts from pursuing – in addition – alternative evaluation approaches if and when they have reason to believe that the alternatives would yield more valid results or might better reflect the needs of the target audience of an analysis.

### **Washington Panel**

A group of experts known as the *Washington Panel* was convened by the U.S. Public Health Service with the main task to develop standards for cost-effectiveness analysis (CEA), in order to ensure that differences in reported health outcomes, costs, and cost-effectiveness ratios across studies and interventions reflect true differences in the consequences, as opposed to artifacts due to unnecessary differences in method. Within the field of health economics, the Washington Panel introduced the notion of a reference case in 1996. The Panel endorsed the use of CEA as an aid to, not a complete procedure for, decision-making, on grounds of its broader acceptance among health care policy makers compared to cost-benefit analysis (CBA) in light of sensibilities that a willingness to pay measure may inherently favor the wealthy over the poor. The Panel recommended adopting a broad societal perspective, considering all changes in resource use and health effects due to an intervention, using a time horizon long enough to capture all relevant

future effects, applying a discount rate of 3% for both costs and effects, and expressing health-related outcomes as Quality-Adjusted Life Years (QALYs). The Panel proposed reporting incremental cost effectiveness ratios (ICERs) but did not suggest an ICER threshold separating cost-effective technologies from others. The convention to exclude “indirect” productivity loss from cost calculation for reference case analysis, introduced by the Washington Panel for concerns about double-counting (assuming the full impact of morbidity was captured in the QALY measure and hence part of the denominator of the ICER), became subject to controversial debate among health economists.

### **National Institute for Health and Clinical Excellence (NICE)**

NICE was established as a Special Health Authority within the United Kingdom National Health Service (NHS) in 1999 that quickly attained high international visibility. NICE evaluates 20-30 (mainly new and mainly pharmaceutical) technologies each year, and provides mandatory guidance on use to the NHS in England and Wales on grounds of their clinical and cost-effectiveness. In order to improve consistency within and between technology appraisals, NICE adopted a generic reference case with its revised methods guide in 2004. NICE justified the focus on CEA using the QALY, assumed to represent a universal and comprehensive measure of health outcomes, by its widespread use.

Costing should be done from the perspective of the NHS and include personal social services (PSS); future costs and benefits should be discounted using an annual rate of 3.5%. Since the 2004 methods guidance, parameter uncertainty should be evaluated using probabilistic sensitivity analysis. NICE indicated a most plausible range of ICERs between £20,000 and £30,000 per QALY gained as a benchmark for judgments about the

cost-effectiveness of an intervention, while recognizing that other factors such as the degree of clinical need of patients may influence its appraisals. According to NICE, estimates of the NHS (and PSS where appropriate) budgetary impact (“affordability”) of adopting a technology are not used for decision-making but for implementation planning only. NICE allows additional (non-reference case) analyses if and when these can be justified.

[TABLE 1 ABOUT HERE]

### **Context and Critique**

The concept of a reference case has not been universally adopted among international decision-making bodies and health technology assessment (HTA) agencies using economic evaluations. For instance, the revised Australian guidelines, issued by the Pharmaceutical Benefits Advisory Committee (PBAC) in November 2006, expressed a general preference for cost-utility analysis (i.e., CEA using health-adjusted life years – most often QALYs – as a measure of health-related outcomes) but explicitly supported the use of CEA (with health outcomes measured in natural units, such as mm Hg blood pressure reduction, episode-free days, clinical events avoided, or [unadjusted] life years gained; however, the choice of outcome measure should be justified) and cost-consequence analysis when disaggregation of outcomes would be helpful. PBAC is also prepared to accept supplementary cost-benefit analysis (CBA) where outcomes are measured in monetary terms. The PBAC guidelines thus provide for an important example where greater flexibility of analytic approaches is endorsed.

This notwithstanding, current international methodological guidelines for health economic evaluations broadly agree on many salient aspects such as type of analysis (CEA), strong reliance on clinical effectiveness data and the principles of evidence-based medicine (Cochrane-style systematic reviews), choice of comparators, incremental comparisons reporting ICERs, the need to address decision uncertainty by way of sensitivity analysis, the need for and acceptance of decision analytic modeling, and adequacy of time horizon. There is less agreement among guidelines on the appropriate perspective of analysis (with a payers' perspective more often recommended in formalized official guidelines, as opposed to a societal perspective in informal academic guidelines), the relevance of phase III efficacy trials, and the role of modeling. Ongoing academic debate concerns the valuation of health outcomes (e.g., natural units versus QALYs versus willingness to pay), how best to account for uncertainty (e.g., regarding the use of probabilistic sensitivity analyses), and the role of budget impact analysis.

A major impetus behind the advocacy of a reference case approach by the Washington Panel and by NICE has been the basic ability to rank technologies across different disorders by their incremental cost per QALY, and therefore the assumption that such rankings ("league tables") are conceptually valid. The implicit normative premises, in particular the value judgment of a primary health service objective to maximize the distribution-independent sum of QALYs produced (given a budget constraint), are not universally shared and have been described as empirically flawed, i.e., not reflecting prevailing public preferences. Accordingly both the Washington Panel and NICE have

acknowledged the need to consider other factors beyond those specified for reference case analysis. Some observers have noted that, in practice, adherence to a generic standard may contribute to a neglect of disease-specific information and thus contradict the aim to use the best available clinical evidence in the context of HTAs. Also concerns have been raised that high levels of standardization might foster analyses ‘by the cookbook’ and thwart further methods development. However, the usefulness of the reference case approach is perhaps best demonstrated by the fact that the absence of a methodological standard, and therefore inconsistency of methods applied, has been cited as a reason why CBA (using contingent valuation to establish the willingness to pay for health care interventions) has not yet had much policy impact – despite its theoretical advantages and a growing number of published CBAs.

*Michael Schlander*

*See also* Contingent Valuation, Cost Measurement Methods, Cost-Benefit Analyses, Cost-Effectiveness Analysis, Cost-Utility Analysis, Discounting, Health Status Measurement, Pharmacoeconomics, Quality-Adjusted Life Years (QALYs), Technology Assessments, Willingness to Pay

### Further Readings

Australian Government, Department of Health and Ageing (2006). *Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee*. Canberra, ACT: Pharmaceutical Benefits Advisory Committee.

Dolan, P., Shaw, R., Tsuchiya, A., & Williams, A. (2005) QALY maximisation and people's preferences: a methodological review of the literature. *Health Economics* 14 (2): 197-208.

Gold, M.R., Siegel, J.E., Russell, L.B., & Weinstein, M.C. [*Washington Panel*] (Eds.). (1996). *Cost-Effectiveness in Health and Medicine*. New York, NY, & Oxford: Oxford University Press, especially Appendix A, pp. 304-311.

Hjelmgren, J., Bergren, F., & Andersson, F. (2001). Health Economic Guidelines – Similarities, Differences and Some Implications. *Value in Health* 4 (3): 225-250.

National Institute for Clinical Excellence (2004). *Guide to the Methods of Technology Appraisal*. London: NICE.

Richardson, J. (1994). Cost Utility Analysis: What Should be Measured? *Social Science & Medicine* 39 (1): 7-21.

Sach, T.H., Smith, R.D., & Whynes, D.K. (2007). A 'League Table' of Contingent Valuation Results for Pharmaceutical Interventions – A Hard Pill to Swallow? *Pharmacoeconomics* 25 (2): 107-127.

Schlander, M. (2007). *Health Technology Assessments by the National Institute for Health and Clinical Excellence: A Qualitative Study*. New York, NY: Springer.

**Table 1: Overview of reference case definitions**

For comparison, methodological guidelines may be informal (“i”; usually academic) or formalized (“f”; issued by official bodies such as HTA or pricing and reimbursement agencies)

<b>Issue</b>	<b>Washington Panel reference case</b>	<b>NICE reference case</b>	<b>Methodological guidelines</b>
Problem definition	The Panel’s framing recommendations are kept separate from its reference case definition	Scope from NICE	Usually expected to define indication, patient (sub)groups, comparator, and perspective
Comparator(s)	Existing practice; if not cost-effective, consider a (a) best available, (b) viable low cost, or (c) “do nothing” alternative	Alternative therapies routinely used within the NHS; will be defined in the scope developed by NICE and will require definition and justification	Usually common practice (“f”); however, somewhat vague (“existing practice”, “common practice”)
Evidence on outcomes	Data should be selected from the best designed (and least biased) sources that are relevant to the question and population under study	Systematic review, with a preference for quantitative meta-analysis of randomized clinical trials data	Usually (long-term) effectiveness, not efficacy; with a broadly prevailing preference for data from randomized clinical trials
Economic evaluation	Cost-effectiveness analysis (CEA)	Cost-effectiveness analysis (CEA)	Usually cost-effectiveness analysis (CEA); sometimes more flexible (including cost-minimization and cost-benefit analysis, CBA)
Perspective on outcomes	All health effects, encompassing the range of groups of people affected, over a time horizon long enough to capture all relevant future effects	All direct health effects on individuals, whether patients or others (principally caregivers); time horizon should be sufficiently long to reflect any differences between the technologies being compared	Usually all relevant health outcomes
Perspective on costs	Societal perspective, long-term using opportunity cost; excluding indirect (productivity) costs; perspective should be explicitly identified	National Health Service (NHS) and personal social services (PSS)	Heterogeneous; direct health care costs only or direct and indirect (productivity) costs (“f”); societal perspective requested more often in informal guidelines (“i”)

			guidelines (“i”)
Discount rate	A real, riskless discount rate of 3.0% should be used, complemented by sensitivity analysis (drawn from 0% to 7%, including 5%)	An annual rate of 3.5% p.a. on both costs and health effects	Often 5% discount rate (“f”); heterogeneous recommendations from 2.5% to 10% in informal guidelines (“i”)
Addressing uncertainty	Univariate sensitivity analysis as a minimum; multivariate sensitivity analyses recommended	Probabilistic sensitivity analysis mandatory (or, where appropriate, stochastic analysis of patient-level data)	Sensitivity analysis
Measure of health benefits	Quality-adjusted life years (QALYs)	Quality-adjusted life years (QALYs)	Usually including QALYs, with more flexibility as to other measures (“f”, “i”), especially physical units; sometimes willingness to pay
Source of preference data for calculation of utility weights	Community preferences; if unavailable, patient preferences may be used as an approximation	Representative sample of the public (UK)	If QALYs are used, usually community preferences
Health state valuation method	Quality weights must be preference-based and interval-scaled	Choice-based method (for example, time trade-off or standard gamble; not rating scale)	If QALYs are used, usually choice-based method; often standard gamble and time-trade off; sometimes rating scales (!)
Description of health states for calculating QALYs	A generic classification scheme, or one that is capable of being compared to a generic system	Using a standardized and validated generic instrument	Heterogeneous; sometimes disease-specific instruments allowed (“f”)
Equity position	Discussion of roles and limitations of CEA in Introductory Chapter (separate from reference case definition)	Each additional QALY has equal value	n.a.
Budget impact analysis	n.a.	Impact on NHS not part of the decision-making process; however, required to allow effective national and local financial planning	Usually n.a.; Ontario: products with high budget impact will need more rigorous documentation of cost-effectiveness