Cost-Utility Analysis in Health Care: All models are wrong, but this one is useful

Cassandra Mah, Simon Fraser University

Abstract

This paper was originally written for Dr Stuart Peacock's HSCI 473 course *Valuing Health and Health Outcomes*. The assignment asked students to critically assess the advantages and limitations of cost-utility analysis as a method for evaluating value for money in health care. The paper uses APA citation style.

Introduction

Recent scientific development has led to increased availability of new health technologies including pharmaceuticals, diagnostic tests, and medicals devices which has been accompanied by increased demand. In resource-constrained health systems, an important goal is to maximize some measure of benefit. To support this, economic evaluation is a comparative analysis of alternative courses of action that considers both costs and outcomes. One form of economic evaluation is cost-utility analysis (CUA), which measures outcomes in terms of preference-based quality of life (QOL). While CUA is required by many health technology assessment (HTA) agencies, CUA is limited by its data sources and assumptions, quality of life measurement, and a particular disregard for equity. However, these limitations are not inherent to CUA, and there are research programmes underway to address them. CUA's generic and preference-based outcomes make it a critical component of assessing value for money in resource allocation.

What is Cost-Utility Analysis?

CUA can be conducted through model-based or trial-based evaluation involving multiple components. The key difference between CUA and other forms of economic evaluation is that CUA measures outcomes in terms of

quality-adjusted life years (QALYs). QALYs simultaneously incorporate length of life and quality of life into a single unit ranging from 0 (dead) to 1 (full health) (Drummond, Sculpher, Claxton, Stoddart, & Torrance, 2015). To calculate QALYs, utility weights for health states can be derived directly, with individuals valuing their own health status by completing exercises such as time-trade off or standard gamble, but QALYs are more commonly derived indirectly, where individuals complete a questionnaire to describe their health state, which corresponds to an 'off the shelf' utility weight. This utility weight is multiplied by the time spent in that health state and change in QALYs over the duration of a treatment is calculated. Next, the treatment of interest is compared to the standard of care by dividing the difference in costs by the difference in effects (QALYs) to derive an incremental cost-effectiveness ratio (ICER). The ICER can be compared against implicit or explicit thresholds to help inform reimbursement decisions by policymakers. CUA is advantageous for informing policymaking for two core reasons.

Advantages of CUA: Generic and Preference-Based Outcomes

One of the foremost benefits of CUA that its outcomes (QALYs) are generic in nature, which facilitates comparison across disease areas. While clinical effectiveness is commonly considered in funding decisions (Cromwell, Peacock, & Mitton, 2015), given that it is measured using a variety of outcomes, it becomes difficult to compare results. For example, comparing cost per depression-free day against cost per knee surgery is challenging because the outcomes are measured in different units. Common instruments used to derive QALYs are designed to capture elements of health-related QOL (HRQoL) universal to all conditions, thereby standardizing outcomes. Compared to other generic outcomes such as mortality, QALYs are superior because with the increase of chronic conditions, health care is no longer solely focused on saving lives but also improving their quality. QALYs are still more informative than a generic disability-adjusted life year (DALY) because while DALYs may enable comparison across disease areas and account for disability, they do not represent preferences, meaning we cannot know the relative amount of value people have for different health states.

Another core advantage of QALYs in CUA is that they represent the preference that people have for different health states, which is useful for funding decisions with limited budgets. While it may seem reasonable to assume that people value life-saving interventions more than minor cosmetic procedures, ascertaining precise estimates of what people are willing to sacrifice (i.e.

opportunity cost: the benefit foregone from pursuing one course of action) to obtain that health status is important because it reflects the funding decision context, where a gain for some will inevitable result in a loss for others. This model not only informs resource allocation across types of diseases, but also indicates what treatment for a given health condition is likely to provide the best value for patients. Consequently, QALYs move beyond the traditional assumption that improvements in conventional clinical outcomes are always proportionate to improvements in quality of life. Additionally, as preference-based units, QALYs are favourable because they approximate utility for health states *without* relying on monetary outcomes. Another form of economic evaluation called cost-benefit analysis measures an intervention's utility as the amount people are willing to pay for it. Not only is it challenging to elicit willingness to pay in contexts with publicly funded health care where individuals have little reference for their personal thresholds (Brazier, Ratcliffe, Saloman, & Tsuchiya, 2017), but willingness to pay is effectively a function of individuals' *ability* to pay. This may have equity implications that conflict with the 'preferences' people have for equity. Thus, the comparability and preference-based nature of QALYs make CUA an essential piece of evidence in considering whether to fund a treatment. Nonetheless, CUA has several limitations.

Limitations of CUA

The challenges of CUA and efforts to address them are discussed in three broad themes: general methods, QALY measurement, and disregard for equity.

General Methodological Challenges

A CUA is limited by the quality of its data and by certain modelling assumptions, though effects vary with evaluation type. Trial-based evaluations based on single studies are susceptible to issues of study design. Potential selection bias may reduce exchangeability between intervention groups, compromising internal validity. Inclusion/exclusion criteria for trials may restrict the study sample to contain unrealistic distributions of characteristics, thereby limiting external validity. Trial follow-up may result in time horizons insufficient for capturing an intervention's long term effects. Certain trials do not collect HRQoL data, with over half of CUAs submitted to the pan-Canadian Oncology Drug Review relied on HRQoL data from *other* studies, even studies of indications *other than* that under evaluation (Raymakers, Regier, & Peacock, 2018). Where trial-based evaluations are not ideal, model-based evaluations such as decision trees or

Markov models can be conducted that rely on diverse forms of evidence. Model-based CUAs can address the previously discussed inadequacies by comparing more treatment options, reflect more evidence, link intermediate to final end points, extrapolate longer time horizons to capture effects, customizing results to the decision context, and assessing heterogeneity in the results (Drummond et al., 2015). The synthesis of diverse evidence sources still depends on data quality and models are subject to assumptions and judgments that should be carefully scrutinized. In addition to these more general methodological challenges, CUA is often criticized for using QALYs.

Critiques of QALYs as Operationalized Currently

While the use of QALYs are indeed CUA's major advantage, the current operationalization of the QALY model has flaws: QALYs may not adequately capture full treatment benefit, health state descriptions lead to different valuations, and different preference-based instruments produce inconsistencies in index scores. Given that instruments vary widely in length, content coverage of the descriptive system, and valuation method, variation in index scores occur (Richardson, Iezzi, & Khan, 2015). For instance, instruments such as the Health Utilities Index Mark 3 focus narrowly on physical health, and consequently they perform poorly for capturing elements of QOL for people with mental health conditions, as compared to instruments such as the Assessment of Quality of Life 8 Dimension that describe health more comprehensively (Engel, Chen, Richardson, & Mihalopoulos, 2018). Therefore, choice of instrument to derive QALYs is critical, despite the fact that a primary benefit of QALYs is that they ought to permit comparability. With regard to accusations that the QALY model discriminates against disability, Whitehurst and Engel demonstrate that the basis of disagreement between public and patient values lies in the health state *description* valued by the public rather than a systematic failure to capture the worth of individuals with disability (2018). Further, QALYs imply a focus on outcomes that does not account for process utility, although there is growing evidence of personal utility of the value of diagnostic or genetic information irrespective of the test's impact of health outcomes (Regier, Weymann, Buchanan, Marshall, & Wordsworth, 2018).

Several programmes of research have attempted to address concerns regarding QALY scope. The 'Extending the QALY' project is developing a new instrument to capture broad benefits in health, social care, and for carers (E-QALY Team, n.d.). Another alternative is the Capability Approach, which

conceptualizes QOL as the *ability* to be and do certain things rather than measuring people's functionings. This approach is operationalized by the collection of ICECAP and Adult Social Care Outcomes Toolkit measures, which derive preference-based 'capability-adjusted life-years' that could be outcomes in CUA (Al-Janabi, Flynn, & Coast, 2012; Van Loon, Van Leeuwen, Ostelo, Bosmans, & Widdershoven, 2018). While the Capability Approach may be grounded in equity-related principles, equity is disregarded in conventional CUA.

CUA Disregards Equity

Reducing inequalities is often an important part of health care systems, yet CUA (and economic evaluation generally) does not systematically evaluate equity. CUA examines "average" marginal effects of treatments and does not consider its distributional effects. Many healthcare systems have mandates to reduce inequalities in health and equity ("reducing inequalities between groups") is a frequently used criteria among published priority setting exercises (Cromwell et al., 2015). Nevertheless, definitions of equity or thresholds for tolerable inequality are often unclear whether they refer to vertical equity, which holds that people with greater health need should receive more health care than those with less need, or horizontal equity, which implies that those with the same need receive the same care) and inconsistent across settings. Hence, Johri and Norheim's review recommends that HTA bodies enlist techniques for explicit consideration of equity to improve overall decisions and procedural performance (2012). Given that funding and policies are not neutral, CUA may perpetuate, exacerbate, or improve inequalities, yet there is no accepted mechanism to examine this systematically. Although CUA conventionally disregards equity, it remains susceptible to change.

Moving Forward on Equity Challenges

Several bodies of work have developed formal methods to incorporate equity into cost-effectiveness analysis, which can be broadly conceptualized as either equity impact analysis or equity trade-off analysis. Equity impact analysis quantifies the distribution of costs and effects by "equity-relevant" variables such as socioeconomic status or severity of illness. These variables are selected based on their relevance to the decision problem and whether the policymakers and/or public are adverse to inequalities due to said variables. CUA could either take the form of an extended CUA that analyzes the 'distribution of both health benefits and financial risk protection benefits per dollar expenditure', or a distributional

CUA which focuses on the distribution of health effects and health opportunity costs from a limited health care budget, and aggregates all costs/effects into a summary metric to be compared against a measure of health equity {Formatting Citation}. Both forms of health equity impact analysis makes the distribution between groups explicit, supporting the transparency pillar in Daniels and Sabin's Accountability for Reasonableness framework for health care (1998). To go further, two main forms of equity trade-off analysis quantify the trade-offs between improving total health and other equity objectives. Firstly, equity constraint analysis effectively counts the "cost" of choosing a fairer but less cost-effective option. Secondly, equity-weighting analysis uses weights applied to groups with various characteristics (i.e. severity of illness or income inequalities), that can be applied to QALYs in CUA (Cookson et al., 2017). As these methods are new, potential for future work lies ahead.

Challenges and Opportunities for Equity

This paper has demonstrated that equity-informed CUA is feasible, though significant challenges remain. Firstly, incorporating equity in CUA requires additional data on the social distributions of parameters for equity variables of interest (Cookson et al., 2017). Secondly, a challenge to using the equity-informed CUA is a lack of accepted sources for quantified trade-offs in equity from social welfare functions (Cookson et al., 2017). While studies have found that the public in commonwealth countries are generally adverse to health inequalities, the degree to which people value different types of inequalities such as general health inequalities versus inequalities based on social variables that are inherited (i.e. ethnicity) or debatably behavioural (i.e. income) is inconclusive (Blacksher, Rigby, & Espey, 2019; Mcnamara, Holmes, Stevely, & Tsuchiya, 2019; Norman, Hall, Street, & Viney, 2013). Further, methods to derive inequality aversion differ (Mcnamara et al., 2019), which could explain variation in preferences and limit comparability. So long as HTA bodies do not define a reference level of health inequality, distributional CUA may be most suitable for considering equity until there are further developments in equity weighting.

Conclusion

Cost-utility analysis remains an essential tool to support value- and evidence-based policymaking. There are significant methodological challenges to conducting CUA, which broadly concern modelling assumptions (data availability/quality, model perspective, time horizon), measuring QALYs

(instrument description and valuation), and equity considerations. That CUA outcomes are generic and preference-based are important advantages, and so future work should further work should pay attention to developing.

References

Al-Janabi, H., Flynn, T. N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. *Quality of Life Research*, *21*(1), 167–176. https://doi.org/10.1007/s11136-011-9927-2

Blacksher, E., Rigby, E., & Espey, C. (2019). Public Values, Health Inequality , and Alternative Notions of a "Fair" Response, *35*(6), 889–920. https://doi.org/10.1215/03616878-2010-033

Brazier, J. E., Ratcliffe, J., Saloman, J., & Tsuchiya, A. (2017). *Measuring and Valuing Health Benefits for Economic Evaluation* (Second). Oxford, UK: Oxford University Press.

Cookson, R., Mirelman, A. J., Grif, S., Asaria, M., Dawkins, B., Norheim, O. F., … Culyer, A. J. (2017). Using Cost-Effectiveness Analysis to Address Health Equity Concerns. *Value in Health*, *2017*(20), 206–212. https://doi.org/10.1016/j.jval.2016.11.027

Cromwell, I., Peacock, S. J., & Mitton, C. (2015). 'Real-world ' health care priority setting using explicit decision criteria: a systematic review of the literature. *BMC Health Services Research*, 1–11. https://doi.org/10.1186/s12913-015-0814-3

Daniels, N., & Sabin, J. (1998). The Ethics of Accountability in Managed Care Reform. *Health Affairs*, *17*(5), 50–64. https://doi.org/10.1377/hlthaff.17.5.50

Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, L. G., & Torrance, G. W. (2015). *Methods for the Economic Evaluation of Health Care Programmes* (4th ed). Oxford, UK: Oxford University Press.

Engel, L., Chen, G., Richardson, J., & Mihalopoulos, C. (2018). The impact of depression on health-related quality of life and wellbeing: identifying important dimensions and assessing their inclusion in multi-attribute utility instruments. *Quality of Life Research*, *27*(11), 2873–2884. https://doi.org/10.1007/s11136-018-1936-y

Johri, M., & Norheim, O. F. (2012). Can Cost-Effectiveness Analysis Integrate Concerns For Equity? Systematic Review, *28*(2), 125–132. https://doi.org/10.1017/S0266462312000050

Mcnamara, S., Holmes, J., Stevely, A. K., & Tsuchiya, A. (2019). How averse are the UK general public to inequalities in health between socioeconomic groups ? A systematic review. *The European Journal of Health Economics*, 1–11. https://doi.org/10.1007/s10198-019-01126-2

Norman, R., Hall, J., Street, D., & Viney, R. (2013). Efficiency and Equity: A Stated Preference Approach. *Health Econ*, *581*(April 2012), 568–581. https://doi.org/10.1002/hec

Raymakers, A. J. N., Regier, D., & Peacock, S. (2018). Inclusion of health-related quality-of-life data in oncology drug reimbursement submissions in Canada: A review of submissions to the pan-Canadian Oncology Drug Review. *Journal of Clinical Oncology*, *36*(30_suppl), 94. https://doi.org/10.1200/JCO.2018.36.30_suppl.94

Regier, D. A., Weymann, D., Buchanan, J., Marshall, D. A., & Wordsworth, S. (2018). Assesing the Value of Next-Generation Sequencing Valuation of Health and Nonhealth Outcomes from Next-Generation Sequencing: Approaches, Challenges, and Solutions. *Value in Health*, *21*(9), 1043–1047. https://doi.org/10.1016/j.jval.2018.06.010

Richardson, J., Iezzi, A., & Khan, M. A. (2015). Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. *Quality of Life Research*, *24*(8), 2045–2053. https://doi.org/10.1007/s11136-015-0926-6

Team, E.-Q. (n.d.). Extending the QALY. Retrieved November 10, 2019, from https://scharr.dept.shef.ac.uk/e-qaly/

Van Loon, M. ., Van Leeuwen, K. M., Ostelo, R. W., Bosmans, J. G., & Widdershoven, G. . (2018). Quality of life in a broader perspective: Does ASCOT reflect the capability approach ? *Quality of Life Research*, *27*(5), 1181–1189. https://doi.org/10.1007/s11136-017-1756-5

Whitehurst, D. G. T., & Engel, L. (2018). Disability discrimination and misdirected criticism of the quality- adjusted life year framework, *44*(11), 793–795.

By submitting this essay, I attest that it is my own work, completed in accordance with University regulations. I also give permission for the Student Learning Commons to publish all or part of my essay as an example of good writing in a particular course or discipline, or to provide models of specific writing techniques for use in teaching. This permission applies whether or not I win a prize, and includes publication on the Simon Fraser University website or in the SLC Writing Commons Open Journal.

Available from: https://journals.lib.sfu.ca/index.php/slc-uwc

SFU LIBRARY DIGITAL PUBLISHING