

Health Valuation: Demonstrating the Value of Health and Lifespan

Elly A. Stolk¹ · Benjamin M. Craig² · Brendan Mulhern³ · Derek S. Brown⁴

Published online: 8 June 2017
© Springer International Publishing Switzerland 2017

1 Introduction

In health preference research, some studies attempt to mimic a specific clinical decision (e.g., choosing between surgery and radiation), while others focus on understanding the value of health and lifespan by exploring their tradeoffs [1]. The latter approach, known as health valuation, summarizes evidence on health and lifespan to facilitate a wide range of policy analyses, including burden-of-illness studies and cost-effectiveness analyses (CEAs). Traditional approaches to health valuation use complex adaptive tasks, equating gains in health to increases in lifespan (time trade-off) or decreases in risk of death (standard gamble). More recently, non-adaptive tasks, such as paired comparisons, have been identified as a promising alternative to adaptive tasks. In this paper, we review the motivations for health valuation, discuss common concepts such as quality-adjusted life years (QALYs), and summarize the next steps in this emerging field.

2 Why Conduct a Health Valuation Study?

In many countries, a formal health technology assessment (HTA) is required to inform access or coverage decisions relating to new or existing technologies [2]. HTA agencies, such as the National Institute for Health and Care Excellence (NICE) in the United Kingdom, consider causal evidence on benefits and costs as well as the relative value of health interventions from multiple perspectives (e.g., societal, patient, etc.) when providing guidance and advice for improving health and social care.

Although discrete choice experiments (DCE) between health-related goods and services (e.g., radiation vs surgery) can assess patients' preferences regarding the costs and benefits of alternatives within a therapeutic class, there is also a need to separate the assessment of the benefits (through societal or patient perspectives) from the cost implications (through payer perspectives), particularly when the beneficiary does not pay the full cost of the intervention. Given the available budget, HTA agencies regularly weigh the interests of multiple stakeholders and conclude whether an intervention is 'good value for money'. These appraisals are greatly facilitated by using standardized methods for describing and valuing health, so that effects can be compared across disease areas and related to cost.

A preferred strategy for measuring health outcomes in applied research is to use standardized instruments that ask respondents to classify their health on a fixed set of dimensions (i.e., patient-reported outcomes [PRO]). The term 'health valuation' is reserved for the scientific process that elicits preferences for the possible outcomes within a descriptive system. This preference evidence produces values that can be applied to PRO responses creating a preference-based measure that summarizes the evidence on

✉ Elly A. Stolk
stolk@euroqol.org

¹ EuroQol Office, Marten Meesweg 107, 3068 AV Rotterdam, The Netherlands

² Department of Economics, University of South Florida, 4202 East Fowler Avenue, Tampa, FL, USA

³ CHERE—Centre for Health Economics and Research Evaluation, University of Technology Sydney, PO Box 123 Broadway, Sydney, NSW 2007, Australia

⁴ Washington University in St. Louis, Brown School, Campus Box 1196, One Brookings Drive, St. Louis, Missouri 63130, USA

health outcomes. Such measures are currently in high demand for use in HTA, but also in routine clinical practice in, for example, the UK National Health Service.

Health valuation studies have not been conducted for all PRO instruments. Those with preference-based values are typically brief, describe a wide range of health outcomes (including mild and severe outcomes), and have been validated across clinical areas. In complement to generic PRO instruments with these properties such as the EQ-5D, some disease-specific counterparts are available. The merit of a preference-based measure should be evaluated by the psychometric properties of the PRO instrument, the ability of the descriptive system to classify PRO responses and the validity of the preference-based values.

3 Quality-Adjusted Life Years (QALYs)

Most health valuation tasks investigate tradeoffs between health problems that may be purposefully chosen to be relevant for a wide range of conditions and thus enable comparisons of health across patient groups and treatments. Due to the breadth of outcomes, it is particularly useful to summarize their value using a common metric (similar to a monetary unit; e.g., the US dollar). A QALY is a year with no health problems, specifically extending a person's lifespan from 'immediate death' to an idealized lifespan of 1 year in full health. Since all treatments aim to affect either health or lifespan or both, a combined metric offers the greatest scope for comparison across patient groups and treatments. This wide scope for application makes the QALY the universal numéraire for health valuation studies.

Understanding values for health outcomes was the primary purpose of introducing the QALY, but the concept does not prescribe what kind of outcomes ought to be valued or whose perspective counts. Some values are based on a generic descriptive system (e.g., the EQ-5D), while others focus on the domains of specific conditions such as dementia [3]. Values may represent either societal preferences or the preferences of a patient or subpopulation (e.g., women) [1]. Overall, HTA agencies favor values based on societal preferences for decisions regarding the allocation of public resources because this allows them to summarize the benefits from the perspective of the taxpayer.

Expressing value of health outcomes on a QALY scale has some well known limitations [2–4]. First, the QALY scale may poorly summarize the impact of a disease on a family or other communal unit. Further issues that arise involve individual preferences over the health of others and how to account for childhood development, pregnancy, and reproductive health. Similarly, the QALY scale does not come with a specific suggestion for how to measure value except that the technique be preference-based. Yet, the

sensitivity of preference-based measures is dependent on the psychometric properties of the PRO instrument, the uncertainty in the preference-based values, and the experimental and analytical methods used to derive these values [5]. In spite of user satisfaction with the current state of play, amongst scientists, the strategy for health valuation remains subject to debate and innovation.

4 Adaptive and Non-Adaptive Tasks in Health Valuation

Systemic differences between valuation tasks may be traced to known biases and limitations. A key characteristic of traditional valuation tasks, such as standard gamble and time trade-off, is the adaptive nature of the tasks. These tasks offer respondents a series of discrete choices along a decision tree in order to identify their points of indifference between health outcomes. The advantage is that respondents report cardinal values ($A = B$), while the approach of gradually homing in on preferences allows them to advance towards a conclusion about this rather than being asked directly. Non-adaptive tasks can only capture preference orderings ($A > B$) and infer value from ordinal responses. Ample proof of principle has been obtained to demonstrate the feasibility of adaptive and non-adaptive tasks for health valuation, and agreement across tasks (e.g., time trade-off tasks typically produce a large gap near full health for slight problems, such as wearing glasses).

The merits of non-adaptive tasks, such as paired comparisons, for health valuation are widely recognized: they are easier for respondents and potentially reflect more closely how people think about health, since it could be questioned whether respondents know their own preferences well enough to answer trade-off questions without making mistakes [5]. Moreover, non-adaptive tasks avoid problems and bias associated with adaptive task design following from their dependence on interviewers, possibility for task shortcutting, and the lack of precision of obtained responses. Such influences imply that adaptive responses depend in part on the respondent's preferences but also on the iterative process and task design, which non-adaptive tasks avoid.

Underlying the design of adaptive tasks is the simplifying assumption: each additional year with no health problems has the same value. Preference evidence casts doubts on this constant proportionality assumption [6]. For example, increasing a quality-adjusted lifespan from 1 to 2 years has a higher value than increasing a quality-adjusted lifespan from 9 to 10 years. Gains in lifespan have a diminishing marginal utility. Likewise, the marginal utility of living with reduced health may increase with duration (e.g., adaptation). In another situation, immediate death

may be preferred over additional years with reduced health (i.e., maximum endurable time). Health valuation is expanding our understanding of time preferences; whether or not these non-linear relationships will be incorporated into the QALY concept, however, remains to be seen.

Even aside from the methodological debate on adaptive tasks, investigators are grappling with how best to incorporate timing, duration, lifespan, and age into health valuation, particularly to improve preference-based values for application in resource allocation decisions [2–4, 6–12]. Respondents may have time preferences that differ greatly from the perspectives of governments or other payers. The issue with time preferences was well known when adaptive tasks were first introduced, but the work was postponed because it was thought to over-complicate the elicitation process. The far simpler data collection procedures of non-adaptive tasks potentially allows this issue to be dealt with, which contributes to interest in further exploring their potential to perhaps at some point replace traditional tasks. Nevertheless, traditional adaptive tasks remain in use due to the policy requirement of comparability across studies and ongoing deliberation about the merits of alternative methodological approaches.

5 Next Steps in Health Valuation

While HTA spurred the field of health valuation initially, the resulting values on a QALY scale are increasingly used to inform a range of other decisions. Cutting-edge studies have focused on changing clinical practice via novel improvements in health valuation tasks, such as experience-based tasks, goal setting, and advanced care planning and directives [13]. Experience-based tasks ask patients to prioritize the health problems that they are experiencing and could, for instance, guide treatment by investigating which health problem to relieve first (i.e., triage). This approach avoids difficulties with perception and enhances clinical relevance, in that it emphasizes how patient preferences should be used to personalize their care. Goal setting is a form of shared decision making (SDM) in which patient priorities guide the allocation of resources toward their goals (e.g., focusing on improving the ability to walk vs achieving a good night's sleep). Advanced care planning and directives are not only about resuscitation (i.e., time trade-off), but can document a person's interest in pain relief during pregnancy, chemotherapy for cancer recurrence, post-Alzheimer's, or organ donation. Each of these involves a more patient-centered form of health valuation.

A better understanding of preferences on health and lifespan can improve the allocation of communal resources (e.g., QALYs from a societal perspective) or the advocacy for an individual's care from her or his own perspective

(e.g., advanced directives). Within health preference research, health valuation is bridging the fields of psychometrics and econometrics [5] and advancing new methods [14]. Measuring health and summarizing its value, however, is not in itself sufficient to improve patients' lives. Preference evidence needs to be actionable. Although further methodological advances are welcome, the future of health valuation will depend on broadening its implementation as a tool for regulators, clinicians, and patients.

Acknowledgements Elly Stolk's contribution to writing up this paper was supported by the EuroQol research foundation [project 2013300]. The authors gratefully acknowledge the administrative support for the preparation of this paper from the International Academy of Health Preference Research (IAHPR). Drs. Craig and Brown serve on the Board of the IAHPR Foundation. Dr. Craig is also a founding co-chair of the Health Preference Research special interest group at the International Society of Quality of Life Research.

Compliance with ethical standards

Conflict of interest Elly Stolk, Benjamin Craig, Brendan Mulhern and Derek Brown declare no conflict of interest.

References

1. Craig BM, Mitchell SA. Examining the value of menopausal symptom relief among US women. *Value Health*. 2016;19(2):158–66.
2. Berg RL. Health Status Indexes: Proceedings of a Conference Conducted by Health Services Research, Tucson, Arizona, October 1–4, 1972. Tucson, Arizona: Hospital Research and Educational Trust.
3. Smith MD, Drummond M, Brixner D. Moving the QALY forward: rationale for change. *Value Health*. 2009;12:S1–4.
4. Edgar A. The ethical QALY : ethical issues in healthcare resource allocations. Haslemere: Euromed Communications; 1998. p. 168.
5. Craig B, Reeve B. Patient-reported outcomes and preference research: igniting the candle at both ends and the middle. *ISPOR Connect Uniting Res Pract*. 2012;18(5):8.
6. Attema AE, Brouwer WB. On the (not so) constant proportional trade-off in TTO. *Qual Life Res*. 2010;19(4):489–97.
7. Carr-Hill RA, Morris J. Current practice in obtaining the “Q” in QALYs: a cautionary note. *BMJ*. 1991;303(6804):699–701.
8. Craig BM. The duration effect: a link between TTO and VAS values. *Health Econ*. 2009;18(2):217–25.
9. Dolan P, Gudex C. Time preference, duration and health state valuations. *Health Econ*. 1995;4(4):289–99.
10. Dolan P, Stalmeier P. The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. *J Health Econ*. 2003;22(3):445–58.
11. Stalmeier PF, et al. On the assessment of preferences for health and duration: maximal endurable time and better than dead preferences. *Med Care*. 2007;45(9):835–41.
12. Sutherland HJ, et al. Attitudes toward quality of survival. The concept of “maximal endurable time”. *Med Decis Mak*. 1982;2(3):299–309.
13. Whitty JA, et al. Assessment of individual patient preferences to inform clinical practice. *The Patient: patient-centered outcomes research*. 2017. doi:10.1007/s40271-017-0254-8.
14. Craig BM, et al. US valuation of health outcomes measured using the PROMIS-29. *Value Health*. 2014;17(8):846–53.