

A NOTE ON QALYS VERSUS HYES

Health States Versus Health Profiles

Paul Dolan
University of Sheffield

Abstract

This paper considers the question of what ought to be valued in the context of measuring the outcomes of healthcare interventions. The answer is discrete health states in the case of the quality-adjusted life-year (QALY) model and an entire health profile in the case of the healthy-years equivalent (HYE) approach. How well the weighted average of values attached to the former approximates the overall value attached to the latter depends on the validity of the assumptions of the QALY model. The paper considers some of the empirical literature relating to them. One of the most important assumptions, which from the limited evidence available appears not to hold, is additive separability. However, it is argued that violation of this assumption does not in itself invalidate the QALY approach, since in some circumstances it might be more appropriate to elicit the value of a health state independently of the states that succeed it. Investigation into this issue is identified as one of the key areas where future research efforts should be directed.

Keywords: QALYs, HYE, Health state valuation, Health profiles

There has been considerable debate in the literature over recent years about the relative merits of the quality-adjusted life-year (QALY) and healthy-years equivalent (HYE) approaches to measuring the benefits associated with different healthcare interventions. Much of the debate has been about how to elicit preferences over health states and profiles, particularly about whether the two-stage standard gamble procedure is formally equivalent to the conventional time trade-off method (9;13). This paper addresses the question of what ought to be valued (rather than how it ought to be valued) and discusses the relative merits of valuing an entire health profile *vis-à-vis* discrete health states. Against this background, it sets out the assumptions of the more restrictive QALY model and looks at the empirical evidence relating to them. The paper then discusses a number of important research questions that arise.

WHAT IS TO BE VALUED?

In very general terms, the answer to this question is a simple one; it is the alternative states of health that an individual experiences over the course of their lifetime. This profile of health clearly consists of two components: the different health states that an individual might be in and the different lengths of time that they might be in each state for. One approach would be to construct profiles for each possible life path and then to elicit preferences over them. This approach has the advantage that it places few restrictions on individual preferences. For example, each individual is allowed to determine for themselves the relative weights they wish to attach to the order or timing of particular health states, or to the effect that different lengths of time in any given state would have on them.

The valuation of profiles of health is the key feature of the HYE approach, which asks individuals to state the number of years in perfect health that are considered equivalent to a particular profile (13). In order for the number of HYE to be a valid representation

of an individual's preferences over different profiles, it is necessary to assume that the individual is risk neutral with respect to (discounted or undiscounted) years in full health (17). Less restrictive still is the ex-ante HYE in which the health profile is framed in terms of uncertainty (i.e., a probability distribution), which only makes the assumption that preferences are monotonic with respect to years of life in full health (5).

However, the problem with the HYE approach is that in most contexts there will be a large number of possible profiles of health, each of which would require preference measurement. Therefore, to allow greater generalizability, an alternative approach would be to elicit preferences for one health state (of a specified duration) at a time. The value of any given profile could then be estimated by taking the (discounted or undiscounted) weighted average of the value for each of the health states in that profile multiplied by the time spent in each state. This is the approach adopted in the calculation of QALYs. Of course, this places greater restrictions on individual preferences, since a number of assumptions have to be made when calculating this weighted average.

THE ASSUMPTIONS OF THE QALY MODEL

Bleichrodt et al. (1) have established the least restrictive conditions under which the QALY model will represent individual preferences over a health profile of constant quality. They show the model will hold if an individual is risk-neutral with respect to gambles over life-years for all health states. Strictly speaking, this requires risk neutrality over undiscounted or discounted years of life, but most authors define and measure risk posture according to the assumption that each year of life yields a constant marginal utility. Therefore, while attitudes toward time are often (and mistakenly) subsumed within attitudes toward risk, the standard QALY model assumes both risk neutrality and neutrality toward the timing of events. In addition, for profiles in which health changes over time, each individual's utility function is required to be strongly separable on the time dimension; that is, the utility derived from a profile of health is equal to the sum of the utility derived from each state in that profile.

For the QALY model to fully represent individual preferences, three main assumptions are required: a) risk neutrality over life-years; b) a zero rate of time preference; and c) additive separability. The empirical evidence relating to each of these assumptions is mixed. With regard to risk attitude, McNeil et al. (11) presented bronchogenic carcinoma patients with a gamble involving a 50% chance of full health for 25 years and a 50% chance of immediate death. The mean certainty equivalent number of years was five, which, assuming no discounting, translates into a risk coefficient of 0.43 (for risk neutrality this figure should be 1), thus indicating moderate risk aversion. Stiggelbout et al. (18) found mild risk aversion in a study of men with testicular cancer; the risk coefficient was 0.74. Verhoef et al. (20) found a similar coefficient (of 0.80) among healthy women, but there was evidence of risk-seeking preferences over gambles involving short durations. Conversely, Mehrez and Gafni (12) found that risk-seeking behavior arose more often when the length of time increased.

There have been some studies that have tried to measure directly individual rates of time preference. On the whole, the results suggest that at the aggregate level the rate of time preference is zero. In a study of economics undergraduates, Cairns (3) found that the timing of a health state did not appear to matter as much as the timing of identical levels of wealth did. Redelmeier and Heller (15), in a study of time preference rates over acute health states, observed discount rates of zero in 62% of the cases. And Dolan and Gudex (4) found that the median discount rate was zero across six states of health. However, all studies have reported wide variation in time preference rates at the individual level.

With regard to additive separability, Richardson et al. (16) asked 63 women to value three discrete breast cancer states and a profile made up of the same three states. The profile value was much lower than the value that would be implied by combining the scores for the discrete states. Kupperman et al. (7) elicited values from 121 women for profiles and states relating to pregnancy outcomes and the remainder of the woman's life. The mean values for the profiles were again lower than that implied by combining the values for the discrete states. The results from both studies suggest that respondents focus more on future health states than on current ones. In the Richardson et al. study, the profile ends with suffering and then death, the knowledge of which "casts a shadow over, or devalues, the enjoyment of earlier life-years." And in the Kupperman et al. study, the valuation for the remainder of the woman's life was the most significant variable in explaining the profile score.

In a different test of additive separability, Krabbe and Bonsel (6) asked 104 students to value two profiles that were identical according to the QALY model: one where a good state was followed by a bad one and one where the bad state was followed by the good one. The results lent some support to the QALY model in that two-thirds of respondents were indifferent to the sequence. Treadwell (19) tested another implication of the additive separability assumption, namely, that if two profiles have the same health state during a given year, then preference between them does not switch if the level of health changes during that year. The results from 185 students suggest that independence is "mostly satisfied."

THE WAY FORWARD

There are many important yet unresolved issues that arise out of the foregoing discussion. With regard to general issues relating to attitudes toward risk and time, there is the need to understand more about the contextual factors that influence both. In addition, the sources of wide variations at the individual level, particularly with regard to time preference, should be the subject of further investigation. Future empirical studies into all the issues raised above should, wherever possible, compare the choices implied by responses to hypothetical questions with the decisions made in direct choices between the same alternatives (10). This has attracted a great deal of attention in the willingness to pay literature (14) and should do likewise in the area of health valuation.

Besides these general issues, there are important questions relating directly to the QALY-HYE debate, particularly to the issue of additive separability. There has been much debate in the literature about the pros and cons of carving up a given health profile into a series of discrete health states, but there has been relatively little investigation into the extent to which the sum of the parts provides a good approximation of the whole. There is some evidence to suggest that differences do exist; specifically, that when separate health events are seen as part of a package, people will dislike profiles of decreasing utility (see reference 8 for evidence of this in nonhealth contexts). However, there is also the need to examine whether the combined value for a series of discrete states and the valuation of an entire profile can be related to one another in any systematic way.

In addition to how the valuation of a profile might differ from that implied by combining the value of discrete states, there is the need to examine why they might differ. It might be that the QALY model is genuinely misrepresenting preferences, but it might also be that respondents are redefining the discrete states that make up the profile so as to incorporate any anxiety that they might feel about their (known) future health status. If this were the case, it would mean that the description of the health state is flawed and not the QALY model itself. As with many of the issues raised here, it might be that useful insights into this issue might be provided by inquiry into the cognitive processes that respondents use to arrive at their responses.

At a normative level, there is also the need for criteria to be established by which a choice between the value of a whole profile and the combined value of different states can be made. For example, if a future health state is not known with certainty, it might be more appropriate to value the previous state independently so that the future state does not contaminate its value. And even if there is knowledge about future health, it might still be considered appropriate to elicit the “goodness” associated with each state in isolation of other states (2). This suggests that in some circumstances the QALY model might be more suitable than the HYE model, even if the sum of the parts does not provide a good approximation of the whole.

All of this raises interesting questions about how health states or profiles should be presented to respondents. It is likely that most people would recognize that their health will deteriorate over time, but most would not allow the value of their current health to be affected by such knowledge (unless they are explicitly asked to, of course). It is remarkable that these central questions have been almost completely ignored as the debate has focused instead on the relative merits of different valuation procedures.

This paper has highlighted some of the important theoretical and empirical questions relating to precisely what should be valued, and it is hoped that future research will address them. Ultimately, though, whether any of these issues matters in a policy sense will be determined by whether resource allocation decisions are affected by using QALYs as opposed to HYEs, and this is perhaps *the* research question.

REFERENCES

1. Bleichrodt H, Wakker P, Johannesson M. Characterizing QALYs by risk neutrality. *Journal of Risk and Uncertainty*. 1997;15:107-114.
2. Broome J. *Weighing goods*. Basil Blackwell; 1991.
3. Cairns J. Future discounting: Health, wealth and time preference. *Project Appraisal*. 1992;7: 31-40.
4. Dolan P, Gudex C. Time preference, duration and health state valuations. *Health Econ*. 1995; 4:289-299.
5. Johannesson M. Quality-adjusted life-years versus healthy-years equivalents: A comment. *J Health Econ*. 1995;14:9-16.
6. Krabbe P, Bonsel G. Sequence effects, health profiles and the QALY model: In search of realistic modelling. *Med Decis Making*. 1998;18:178-186.
7. Kupperman M, Shiboski S, Feeny D, Elkin EP, Washington AE. Can preference scores for discrete states be used to derive preference scores for an entire path of events? *Med Decis Making*. 1997;17:42-55.
8. Loewenstein G, Prelec D. Negative time preference. *Am Econ Rev*. 1991;81:347-352.
9. Loomes G. The myth of the HYE. *J Health Econ*. 1995;14:1-7.
10. Loomes G, McKenzie L. The scope and limitations of QALY measures. *Soc Sci Med*. 1989;28:299-308.
11. McNeil BJ, Weichselbaum R, Pauker SG. Fallacy of the five-year survival in lung cancer. *N Engl J Med*. 1978;299:1397-1401.
12. Mehrez A, Gafni A. An empirical evaluation of two assessment methods for utility measurement for life years. *Socio-Economic Planning Sciences*. 1987;21:371-375.
13. Mehrez A, Gafni A. Quality-adjusted life years, utility theory, and health-years equivalents. *Med Decis Making*. 1989;9:142-149.
14. Mitchell RC, Carson RT. *Using surveys to value public goods: The contingent valuation method*. Resources for the future. Washington, DC; 1989.
15. Redelmeier DA, Heller DN. Time preference in medical decision-making and cost-effectiveness analysis. *Med Decis Making*. 1993;13:212-217.
16. Richardson J, Hall J, Salkfeld G. The measurement of utility in multiphase health states. *Int J Technol Assess Health Care*. 1996;12:151-162.

17. Ried W. QALYs versus HYE: What's right and what's wrong? A review of the controversy. *J Health Econ.* 1998;17:607-625.
18. Stiggelbout AM, Kiebert GM, Kievit J, et al. Utility assessments in cancer patients: Adjustment of time trade-off scores for the utility of life-years and comparison with standard gamble scores. *Med Decis Making.* 1994;14:82-90.
19. Treadwell JR. Tests of preferential independence in the QALY model. *Med Decis Making.* 1998;18:418-428.
20. Verhoef LCG, de Haan AFJ, van Daal WAJ. Risk attitude in gambles with years of life: Empirical support for prospect theory. *Med Decis Making.* 1994;14:194-200.