CURRENT OPINION

Pharmacoeconomics 1998 Feb; 15 (2): 119-127
1170-7690/99/0002-0119/$04.50/0

Valuing Health-Related Quality of Life

Issues and Controversies

Paul Dolan

Sheffield Health Economics Group, and Department of Economics, University of Sheffield, Sheffield, England

Abstract

An important consideration when establishing priorities in healthcare is the likely effect that alternative allocations will have on the health-related quality of life (HR-QOL) of the relevant population. This paper considers some of the important issues and controversies surrounding the valuation of HR-QOL. It considers the theoretical and empirical evidence regarding 3 crucial questions: (i) what is to be valued?; (ii) how is it to be valued?; and (iii) who is to value it? Many important yet unresolved issues emerge and directions for future research are suggested. It is argued that this research agenda should have the gathering and analysis of qualitative data at its forefront.

Traditionally, the impact of healthcare has been measured in terms of its effect on mortality. But, of course, health is much more than merely being alive and there is clearly a need to say something about the health of the majority of people who do not experience premature death. As a result, there now exists many different ways in which morbidity can be defined and measured, and the prevalence and incidence of many conditions and illnesses is well documented. However, it is necessary to not only say what conditions people are experiencing but also to say something about the impact that these conditions have on their lives. This is what the measurement of health-related quality of life (HR-QOL) aims to achieve.

It is important to distinguish HR-QOL from more clinically focused measures which are used as proxies for HR-QOL; for example, the extent to which tumour size, blood pressure or serum cholesterol level is reduced through treatment. The distinction is made clear in a study by Jachuk et al.[1] in which they show that doctors and patients have very different ideas about whether or not antihypertensive treatment results in improvements in health. The doctors (who considered treatment to be successful in all cases) were thinking about whether the patient's blood pressure was reduced or not (and the likely effect that this would have on their future health) whilst the patients (half of whom felt no change in health status and some of whom actually felt worse with treatment) were concerned with the effect that the treatment had on their current HR-QOL.

Whilst the role that individual preferences should play in determining priorities in health and elsewhere is a matter of intense debate amongst economists,[2] there is general agreement within the profession that people's preferences regarding the effectiveness or otherwise of different interventions should have some role.[3] Therefore, health economists look to measures of HR-QOL that are preference-based. Against this background, this article considers the theoretical and empirical issues involved in eliciting valuations for different levels of HR-QOL. Space does not permit a discussion of the issues associated with how these different levels of HR-QOL should be described. Those interested

in these issues are referred to Streiner and Norman[4] and Gold et al.[5]

1. Theoretical Considerations

1.1 What is to be Valued?

In very general terms, the answer to this question is a simple one; it is the alternative types and levels of HR-QOL that an individual might experience during the period of interest. This profile of health clearly consists of 2 components; namely, the different states of health that an individual might be in and the different lengths of time that he/she might be in each state for. One approach would be to construct profiles for each possible path and then to elicit individual preferences over them. The valuation of profiles of health is the key feature of the healthy years equivalent (HYE) approach which asks individuals to state the number of years in perfect health that are considered equivalent to a particular profile.[6] This approach has the advantage that it only places 1 restriction on preferences; namely, that an individual is risk neutral with respect to (discounted or undiscounted) years in full health. But the problem with it is that, in most contexts, there will be a large number of possible profiles of health, each of which would require preference measurement.

To allow greater generalisability, an alternative approach would be to elicit preferences for 1 health state (of a specified duration) at a time. The value of any given profile of health 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 quality-adjusted life-years (QALYs). This places much greater restrictions on individual preferences since a number of assumptions have to be made when calculating this weighted average.[7-9] One of the most restrictive is that the value attached to a particular health state is independent of the state(s) that precede or follow it.

1.2 How is it to be Valued?

If changes in HR-QOL are to be quantified, once a health state descriptive system has been developed, it is necessary to determine what values should be attached to the health states. This raises the question of which valuation method(s) should be used. An important consideration here is the level of measurement that is required. For most purposes, including cost-utility analysis, it is necessary for states to be expressed on an interval scale[10] which provides information on how far apart states are in terms of severity. Typically, the aim is to represent health on a scale whereby death and full health are assigned values of 0 and 1, respectively. The 3 methods that have been widely used to generate valuations with such properties are the visual analogue scale (VAS), the standard gamble (SG) and the time trade-off (TTO), and it is these methods that are the focus of the current paper.

It is worth noting, however, that increasing attention is being paid to alternative techniques which can be used to value the benefits associated with different healthcare interventions. For example, the contingent valuation (CV) approach is being used to elicit the monetary value of alternative interventions whilst conjoint analysis (CA) is being used to infer the relative value attached to the different attributes of an intervention. Good introductions to the use of CV and CA can be found in Donaldson[11] and Ryan and Hughes,[12] respectively.

The VAS requires respondents to rate health states on a scale (typically represented by a vertical 'thermometer-type' line) with 'worst' and 'best' end-points usually represented by 0 and 100, respectively. The SG asks the respondent to choose between the certainty of an intermediate health state and the uncertainty of a treatment with 2 possible outcomes, one of which is better than the certain outcome and one of which is worse. The object is to find the probability attached to the better of the 2 uncertain outcomes at which the respondent is indifferent between the certain and the uncertain alternatives. The TTO asks the respondent to choose

between living for a defined period of time in a poor health state and living for a shorter period of time in full health. The time in full health is varied until the respondent is indifferent between the 2 alternatives. More detailed descriptions of both these methods can be found in Torrance.[13]

Because valuations from the VAS are elicited in a choiceless context and thus do not require people to make trade-offs between different arguments in their utility function, the method is commonly regarded by economists as theoretically inferior to the choice-based SG and TTO methods.

According to the axioms of Expected Utility Theory (EUT), if a utility is expressed as equivalent to a gamble, it is a linear function of the risk involved in the gamble.[14] This has led many to regard the SG as the 'gold standard' for health status measurement.[15] Although there is evidence that people systematically violate the axioms of EUT,[16,17] many still consider the SG to be the most valid method because of the normative appeal of the axioms of EUT.[18] However, EUT focuses only on the expected utility of different outcomes and there is increasing evidence that many people consider this to be an irrational basis on which to make decisions.[19]

The SG is also advocated on the grounds that almost all decisions about healthcare are made under conditions of uncertainty.[20] Whilst this is indeed the case, the appropriateness or otherwise of a valuation method is determined by its ability to act as a proxy for utility and not by its capacity to model the situation being valued.[21] In this respect, the TTO may be considered more appropriate since, by definition, it gives the number of years in full health which are valued equally to a (longer) period in the health profile (or state) being measured. In other words, it collapses the relationship between the health profile (or state), its duration and its value into 1 single measure.[19] However, Dolan and Jones-Lee[22] have shown that for a response to a TTO question to provide a direct and unbiased estimate of health state value, it is necessary that there is no discounting of future utilities.

It is difficult, then, to choose between SG and TTO on theoretical grounds since valuations from neither method can automatically be assumed to map directly onto utility. This is an important point since it implies rejecting the idea that the SG should be regarded as the 'gold standard' for measuring health state values.

Of course, the SG and TTO methods both aim to produce valuations for states of health, so it is worth considering the extent to which, in theory, the methods should produce the same valuations. At one level, since both methods are used to value the same thing (namely, a particular state of health), then the SG and TTO should produce the same results. If this were the case, then much of the heat would be taken out of the debate about which is the most appropriate method. However, based on 2 assumptions about individual preferences, it might be expected that the methods will produce different results. The first is that people are risk averse, implying that they will be reluctant to accept the gamble outcomes in the SG. The second assumption is that people have positive time preference, implying that they will be more willing to give up years of life at the end of a profile, as in the TTO. Thus, both assumptions imply that, for the same health states, SG values will be higher than TTO ones.

1.3 Who is to Value it?

Many consider that it is most appropriate to elicit valuations from those people who are currently experiencing the health states for which values are sought. It is argued that these are the only people who know what it is really like to be in those states whilst those without direct experience of health states cannot accurately predict the impact that the states will have on their HR-QOL. But in reality, the distinction between those with and without experience of illness is very blurred. Even in supposedly 'healthy' populations (for example, the general public), there is a substantial degree of ill health; many currently 'healthy' people have experienced ill health at some time in their lives and many have relatives or close friends who are currently experiencing ill health. Moreover, the extent

to which people are able to imagine the impact that certain health states will have on them is likely to vary according to the likelihood of them experiencing those states; people who expect to experience certain health states are likely to have given some thought to the likely consequences of those states. Therefore, it is perhaps more appropriate to consider 'experience' of ill-health as lying on a continuum rather than as being a dichotomy between the 'experienced' and the 'inexperienced'.

Maintaining the dichotomy for the sake of exposition, there may be a number of reasons why the preferences of patients or 'experienced' respondents may not be considered the most appropriate. Since it is now well established that there is a direct positive link between the time spent in ill health and adaptation to that ill health,[23,24] it is questionable whether such adaptation should be taken into account, particularly when allocating resources which will deal with the treatment of prospective patients. Relatedly, if one of the purposes of the healthcare system is to give reassurance to the general public, then resources should, in part, be allocated so as to reassure the public that treatment is available to alleviate the health states they fear the most, even if this fear is in some way misplaced.[25] Finally, it could be argued that since the general public pay for healthcare, their preferences should be used in the resource allocation process; a view consistent with conventional welfare economic theory.

For these reasons, it is common for the preferences of the whole population to be considered the most relevant when comparing interventions that affect different population subgroups.[5] However, when comparing interventions for the same condition, it might be more appropriate to use the values of the patients experiencing that condition.

2. Empirical Considerations

2.1 Do Valuations for Profiles Differ from Valuations for Discrete States?

There has been some empirical investigation into the extent to which the value attached to a particular health state is independent of the state(s) that precede or follow it. Both Richardson et al.[26] and Kupperman et al.[27] found that the value of a health profile was significantly lower than the value that would be implied by combining the scores for the discrete health states. These findings suggest that respondents focus more on future health states than on current ones. In the study by Richardson et al.,[26] 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 study by Kupperman et al.,[27] the valuation for the remainder of the life was the most significant variable in explaining the profile score.

2.2 Which is the 'Best' Valuation Method?

Empirical assessment of the relative merits of the SG and TTO methods involves considerations of feasibility, reliability and validity. Feasibility means that the method must be capable of being used in practice and must be acceptable to respondents. It would appear that both the SG and the TTO are feasible in that they have both been widely used in practice and most studies have reported high response rates and even higher levels of complete data.[28] However, in a within-respondent comparison of the SG and TTO in a large sample of the UK general population, Dolan et al.[29] found that a variant of the TTO which used a specially designed board produced fewer missing values than the analogous version of the SG and fewer missing values than variants of the methods which used a self-completion booklet.

Reliability refers to the stability of responses when all pertinent conditions remain unchanged. This can be investigated in 2 ways: i) split-test reliability which assesses an individual respondent's consistency when an item is presented more than once within the same administration; and ii) test-retest reliability which assesses the stability of values over short periods of time. Most studies have found little to choose between the 2 methods in terms of reliability. Torrance[30] found the SG and TTO to have similar split-test correlation coeffi-

cients (between 0.80 and 0.90). Reed et al.[31] found the test-retest reliability (as measured by the correlation coefficient) to be higher for the SG (r = 0.82) than for the TTO (r = 0.74) but these figures were not statistically significantly different from one another. Dolan et al.[29] found that the 'board-based' variant of the TTO performed best, producing a correlation coefficient of 0.81 but this was not significantly higher than the correlation coefficient of 0.71 for the self-administered variant of the SG.

Essentially, a method is valid if it accurately reflects the concept it claims to measure. Strictly speaking, this does not apply to HR-QOL since there is no objective way in which HR-QOL can be valued. However, some attempt must be made to address this issue since there is little point in having a feasible and reliable method that is not measuring the right concept(s). The most rigorous way to establish validity is to test construct validity. A construct is a theoretically derived notion of what an instrument is intended to measure. For example, it might be expected that the rankings of health profiles implied by calculating QALYs from SG and TTO valuations would be the same as the direct rankings of the same profiles from the same respondent. In a test of this, Bleichrodt and Johannesson[32] found that the correlation between the 2 rankings was significantly higher for TTO than for SG.

Overall then, there would appear to be little compelling evidence to favour one method over the other. However, if there is no *a priori* commitment to the normative appeal of the axioms of EUT, then the TTO might be the preferred method on the grounds that it appears to yield the more complete data when administered in a general population sample.

2.3 Do Valuations Differ by Method?

It is important to consider whether in practice the SG and TTO produce similar results, or whether, as suggested above, SG values are higher than TTO ones. Most studies to date have found SG values to be higher than TTO ones[30-36] but others have found the opposite relationship.[29,37] Therefore, whilst empirical evidence on the ordinal relationship between SG and TTO values is mixed, most studies have shown that the 2 methods do yield different valuations from the same respondents for identical descriptions of health.

Even if differences between the valuation methods do exist, and a choice between them is difficult to make, it might be that a systematic relationship exists between the methods. As noted above, although the VAS is commonly regarded by economists as theoretically inferior to the SG and TTO methods, it has the practical advantages of being simpler to complete and cheaper to administer than both of the other methods. Therefore, if an algorithm can be found which maps VAS values onto SG and/or TTO ones, then it might be possible to elicit valuations via (cheap and simple) VAS methods and 'convert' them into (theoretically superior) SG and/or TTO values. In a comparison of mean VAS and TTO values, Torrance[30] concluded that '... the 2 techniques exhibit a systematic relationship [that] can be approximated by ... a logarithmic function and a power function.' Since then, a number of authors have used a power function to estimate SG and TTO valuations from VAS ones.[38-40]

However, there are a number of reasons to be cautious about such findings. First, the power coefficients differ across studies (e.g. a VAS score of 0.10 would convert into a TTO score of 0.23 in the study by Stiggelbout et al.[40] and 0.34 in the study by van Busschbach[39]). Second, the analyses were performed on aggregate- rather than individual-level data, thus making the choice between competing models more difficult as well as making inefficient use of the data. Third, the models presented by Torrance[30] did not hold at the individual level (this is confirmed in studies reported in Dolan and Sutton[41] and Bleichrodt and Johannesson.[42] Fourth, van Busschbach[39] found that the power model offered no improvement over a linear one (in a comparison of a number of different models, Dolan and Sutton[41] conclude that linear ones perform best). Given the current empirical evidence, it would seem that VAS valuations cannot be con-

verted into SG or TTO ones with any degree of confidence.

2.4 Do Valuations Differ by Respondent Characteristics?

Most studies suggest that variation among population subgroups is not explained by the different demographic characteristics of respondents (Froberg and Kane[43] provide an extensive review of the literature). There is, however, some evidence which suggests that preferences over states of health are not entirely independent of the respondent's age. Sackett and Torrance[44] found that TTO valuations increased with age suggesting that people become more tolerant of poor health as they get older, possibly through adapting to a general deterioration in health. Dolan et al.[45] found a somewhat different pattern in that, although valuations rose up to about 40 years of age, there was a general decrease from about 40 to 60 years and then a much sharper fall in later years.

There is also evidence to suggest that experience of illness may influence respondents' valuations of health states. For eaxample, Sackett and Torrance[44] reported that patients receiving dialysis treatment at home assigned higher values to kidney dialysis than did the general public. More recently, Dolan[46] found that current health status has an important effect on valuations, with those in poorer health generally giving higher values. The possibility that valuations differ according to illness experience is consistent with the notion that those in poor health successfully compensate for it (as alluded to above). However, this conclusion is slightly tempered by Llewellyn-Thomas et al.[47] who found that respondents' own health status did not influence health state valuations, and by Daly et al.[48] who found that valuations given to menopausal symptoms did not differ across subgroups of women who were divided on the basis of whether they had experienced these symptoms.

3. A Research Agenda

3.1 Valuing Profiles or States

There has been much debate in the literature about the pros and cons of carving up a given health profile into a series discrete health states but there has been remarkably little investigation into the extent to which the sum of the parts provides a good approximation of the whole. Whilst there is some evidence to suggest that differences do exist, there is 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, since the evidence currently available suggests that respondents focus more on future health states than on current ones, it might be more appropriate in some circumstances to value the previous state independently so that the future state does not contaminate its value. Therefore, it is important that criteria are established by which a choice between the value of a whole profile and the combined value of different states could be made if differences between the 2 are observed.

3.2 Valuing Health or Health Gain

Throughout this article, reference has been made to the valuation of health profiles or health states. But ultimately, it is the value of changes in profiles or states that is important if QALYs or HYEs are to be used to inform resource allocation decisions.[49] Of course, if valuations lie on an interval scale, then the value of a move from state i to state j can be calculated as the value given to state j minus the value given to state i. But it is important to show the extent to which this is the case and, to date, such evidence is almost nonexistent. Testing interval scale properties will enable a more informed choice to be made between profiles and discrete states and between the SG and TTO.

3.3 Harmonisation

Previous studies have used many different descriptive systems, valuation methods and sources of values, thus making comparisons across studies almost impossible. Against this background, a US panel on cost effectiveness in health and medicine[5] recommended a 'reference case' for use in cost-effectiveness analyses. They suggested that the health state descriptive system should be generic, that the valuation method should be preference-based and that the source of values should be a representative sample of the general population. To facilitate comparability between studies, all future studies could use the Gold et al.[5] 'reference-case'.

3.4 Qualitative data

Future studies should be much more interactive than those conducted previously. Of course, this will be much more resource intensive per respondent but there should be a willingness (rather than a reluctance) to trade-off quantitative data for the more detailed qualitative data that intensive questioning could generate. This qualitative data should provide insights into the cognitive processes that respondents use in order to arrive at their responses. Many of the studies referred to in this paper have been written in ways typical of economists; namely, to postulate a null hypothesis, to then collect quantitative data that tests the hypothesis and finally to engage in considerable 'post-hoc' theorising when the results, as invariably happens, do not conform with the null hypothesis. Rather than 'second guessing' respondents, the collection of qualitative data 'straight from the horse's mouth' appears a more appropriate strategy in this context.

4. Conclusion

Very few healthcare interventions have no effect on HR-QOL. So when it comes to allocating resources, it is vital that changes in HR-QOL are taken into account. Some readers may have initially been optimistic about our ability to do this but, in light of the arguments developed in this article, with its emphasis on the theoretical and empirical problems associated with valuing health outcomes, may have become increasingly disillusioned with the whole enterprise. This would be unfortunate because facing up to the violations of certain axioms and the many unanswered questions is better than the alternative of disregarding HR-QOL altogether. This negative response would also ignore the considerable methodological advances that have been made in the field, particularly in the last 20 years. Moreover, many of the issues (how health is to be valued, who is to value it and so on) are issues that are faced by any measure of health outcome – it is just that they are made more explicit when valuing HR-QOL.

It is also important to remember that the violation of certain assumptions (for example, in the QALY model) does not mean that the models concerned should necessarily be abandoned (for example, in favour of something like an HYE approach). Most assumptions can only ever be satisfied approximately and thus a judgment will ultimately have to be made about the extent to which the loss of realism (e.g. of more general QALY-type models) are compensated for by their greater tractability (e.g. compared with less general HYE-type approaches).

Finally, the validity of valuations for health profiles or states does not rest on there being a precise answer to the question of how many QALYs or HYEs a particular programme generates. In many cases, it is likely that the use of different values will make no difference to the ordinal conclusions reached about which programme generates more QALYs or HYEs than which. As Lockwood[50] has argued, '... only a very radical scepticism, according to which one could not even, with any confidence, set numerical limits in such comparisons, would have the effect of rendering the QALY approach wholly useless ... such wholesale scepticism would ... be very difficult convincingly to sustain.'

References

1. Jachuk SJ, Brierley H, Jachuk S, et al. The effect of hypertensive drugs on the quality of life. J R Coll Gen Pract 1982; 32 (235): 103-5
2. Ng YK. Welfare economics. London: MacMillan, 1992
3. Johannesson M, Jonsson B, Karlson G. Outcome measurement in economic evaluation. Health Econ 1996; 5: 279-96
4. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press, 1989
5. Gold M, Siegal JE, Russell LB, et al. Cost-effectiveness in health and medicine. Oxford (NY): Oxford University Press, 1996
6. Mehrez A, Gafni A. Quality-adjusted life years: utility theory, and health-years equivalents. Med Decis Making 1989; 9: 142-9
7. Pliskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. Oper Res 1980; 28: 206-44
8. Miyamoto JM, Eraker SA. Parameter estimates for a QALY utility model. Med Decis Making 1985; 5: 191-213
9. Bleichrodt H, Wakker P, Johannesson M. Characterizing QALYs by risk neutrality. J Risk Uncertainty 1997; 15: 107-114
10. Lipscomb J. Value preferences for health: meaning measurement and use in program evaluation. In: Kane RL, Kane RA, editors. Values and long term care. Lexington (MA): Lexington Books, 1982
11. Donaldson C. Willingness to pay for publicly-provided goods: a possible measure of benefit. J Health Econ 1990; 9: 103-18
12. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. Health Econ 1997; 6: 261-73
13. Torrance GW. Measurement of health state utilities for economic appraisal. J Health Econ 1986; 5: 1-30
14. von Neumann J, Morgenstern O. Theory of games and economic behaviour. New York: Wiley, 1953
15. Torrance GW, Feeny D. Utilities and quality-adjusted life years. Int J Technol Assess Health Care 1989; 5: 559-75
16. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, et al. The measurement of patients' values in medicine. Med Decis Making 1982; 2: 449-62
17. Schoemaker PJH. The expected utility model: its variants, purposes, evidence and limitations. J Econ Lit 1982; 20: 529-63
18. Gafni A, Birch. Preferences for outcomes in economic evaluation: an economic approach to addressing economic problems. Soc Sci Med 1995; 40: 767-76
19. Richardson J. Cost-utility analysis: what should be measured? Soc Sci Med 1994; 39 (1): 7-21
20. Mehrez A, Gafni A. The health-years equivalents: how to measure them using the standard gamble approach. Med Decis Making 1991; 11: 140-6
21. Buckingham K, Drummond M. A theoretical and empirical classification of health valuation techniques. Health Economists Study Group (HESG) Conference; 1993 Jul 3-5; Strathclyde
22. Dolan P, Jones-Lee M. The time trade-off: a note on the effect of lifetime reallocation of consumption and discounting. J Health Econ 1997; 16: 731-9
23. Meyerowitz BE. Postmastectomy coping strategies and quality of life. Health Psychol 1983; 2: 117-32
24. Cassileth BR, Lusk EJ, Strouse TB, et al. Psychosocial status in chronic illness: a comparative analysis of six diagnostic groups. N Engl J Med 1984; 311: 506-11
25. Edgar A, Salek S, Shickle D, et al. The ethical QALY: ethical issues in healthcare resource allocations. Haslemere: Euromed Communications, 1998
26. Richardson J, Hall J, Salkfeld G. The measurement of utility in multiphase health states. Int J Technol Assess Health Care 1996; 12: 151-62
27. Kupperman M, Shiboski S, Feeny D, et al. 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
28. Froberg DG, Kane RL. Methodology for measuring health state preferences II: scaling methods. J Clin Epidemiol 1989; 42: 459-71
29. Dolan P, Gudex C, Kind P, et al. Valuing health states: a comparison of methods. J Health Econ 1996; 15: 209-31
30. Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. Socioecon Plann Sci 1976; 10: 129-36
31. Reed WW, Herbers JE, Noel GL. Cholesterol lowering therapy: what patients expect in return. J Gen Intern Med 1993; 8: 591-6
32. Bleichrodt H, Johannesson M. Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs. J Health Econ 1997; 16: 155-75
33. Wolfson AD, Sinclair AJ, Bombardier C, et al. Preference measurements for functional status in stroke patients: interrater and intertechnique comparisons. In: Kane RL, Kane RA, editors. Values and long term care. Lexington (MA): Lexicon Books, 1982
34. Read JL, Quinn RJ, Berrick DM, et al. Preferences for health outcomes: comparison of assessment methods. Med Decis Making 1984; 4 (3): 315-29
35. Stiggelbout AM, Kiebert GM, Kievit J, et al. Utility assessment 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
36. Lenert LA, Cher DJ, Goldstein MK, et al. The effect of search procedures on utility elicitations. Med Decis Making 1998; 18: 76-83
37. Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients well-being and consequent effect on a cost-effectiveness analysis. J Clin Epidemiol 1992; 45 (5): 505-12
38. Loomes G. Disparities between health state measures: is there a rational explanation? In: Gerrard W, editor. The economics of rationality. London: Routledge, 1993
39. van Busschbach J. The validity of QALYs [dissertation]. Rotterdam: Erasmus University, 1994
40. Stiggelbout AM, Eijkemans MJC, Kiebert GM, et al. The 'utility' of the visual analog scale in medical decsion making and technology assessment: is it an alternative to the time trade-off? Int J Technol Assess Health Care 1996; 2: 291-8
41. Dolan P, Sutton M. Mapping visual analogue scale scores onto time trade-off and standard gamble utilities. Soc Sci Med 1997; 44: 1519-30
42. Bleichrodt H, Johannesson M. An experimental test of a theoretical foundation for rating scale valuations. Med Decis Mak 1997; 17: 208-16
43. Froberg DG, Kane RL. Methodology for measuring health state preferences III: population and context effects. J Clin Epidemiol 1989; 42: 585-92

44. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. J Chronic Dis 1978; 31: 697-704
45. Dolan P, Gudex C, Kind P, et al. The time trade-off method: results from a general population study. Health Econ 1996; 5: 141-54
46. Dolan P. The effect of experience of illness on health state valuations. J Clin Epidemiol 1996; 49: 551-64
47. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, et al. Descibing health states: methodologic issues in obtaining values for health states. Med Care 1984; 22: 543-52
48. Daly E, Gray A, Barlow D, et al. Measuring the impact of menopausal symptoms on quality of life. BMJ 1993; 307: 836-40
49. Fitzpatrick R. A pragmatic defence of health status measures. Health Care Anal 1996; 4: 265-72
50. Lockwood M. Quality of life and resource allocation. In: Bell M, Mendus S, editors. Philosophy and medical welfare. Cambridge: Cambridge University Press, 1988

Correspondence and reprints: Dr *Paul Dolan*, School of Health and Related Research, University of Sheffield, 30 Regent Street, Sheffield S1 4DA, England.
E-mail: P.Dolan@sheffield.ac.uk