HEALTH ECONOMICS
Health Econ. **15**: 543–551 (2006)
Published online 3 January 2006 in Wiley InterScience (www.interscience.wiley.com). **DOI:**10.1002/hec.1074

Does the whole equal the sum of the parts? Patient-assigned utility scores for IBS-related health states and profiles

John Brazier[a], Paul Dolan[a], Korina Karampela[b] and Isabel Towers[a,*]
[a] *Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, UK*
[b] *Global Health Outcomes, GlaxoSmithKline, UK*

Summary

The quality-adjusted life year (QALY) assumes that the value of a health state is linearly related to the time spent in it, which implies that the value of a health state is independent of the states which precede or follow it. Irritable bowel syndrome (IBS) is a suitable condition to test this assumption since it is subject to considerable fluctuations over time.

Forty-nine IBS patients were asked to rate their own health using generic measures of health and a condition specific classification. They were then asked to value five IBS states and four profiles using a self-completed version of the standard gamble technique. The implied value of each profile was estimated using the QALY assumption of linearity over time and compared with the direct profile valuations.

The directly elicited profile values suggest that reductions in the duration of IBS symptoms has less of an impact on the value of quality of life than would be implied by the QALY assumption of linearity over time, though the differences were small. There are a number of competing explanations for this finding, including possible sequence effects, quantity effects or time preference, or it might be due to gestalt effects resulting in a neglect of time spent in symptomatic states of health. Copyright © 2005 John Wiley & Sons, Ltd.

Keywords quality-adjusted life years; valuation of health states versus profiles; irritable bowel syndrome

Background

A commonly used measure of benefit in economic evaluations is the number of quality-adjusted life years (QALYs) that an intervention generates. The QALY combines the value of changes in health-related quality-of-life (HRQoL) with information on length of life into a single index number. The demonstration of cost-effectiveness in terms of an incremental cost-per-QALY ratio is becoming important for public funding of new health care interventions in many countries [1–3]. This paper examines the application of a QALY approach to irritable bowel syndrome (IBS).

Functional gastro-intestinal disorders account for a significant proportion of primary care and hospital outpatient visits. IBS is the most common of these disorders. Patients diagnosed with IBS typically present with a wide range of symptoms, and there is growing evidence that they experience a level of HRQoL significantly below that of the general population. An important feature of this condition is that its symptoms tend to fluctuate considerably over relatively short periods of time. This has important implications both for the assessment of HRQoL in IBS patients and for the extent to which the QALY can represent the benefits associated with their treatment.

*Correspondence to: Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK. E-mail: cmp98imt@sheffield.ac.uk

Received 25 July 2003
Accepted 21 October 2005

The use of QALYs to value the benefits of treatment requires a number of quite stringent assumptions about the way people value a profile of health over time [4]. For example, the number of QALYs associated with a particular health profile, consisting of a given health state, Q, and a given number of years, T, is typically calculated for use in incremental cost-effectiveness analysis by multiplying the value of Q by T. Thus, the value of a health state is assumed to be linearly related to the time spent in that state. This QT QALY model assumes a zero time preference. For health profiles characterised by changes in HRQoL, it is also usually assumed in QALY calculations that the value attached to a particular health state is independent of the state(s) that precede or follow it. This is equivalent to the estimation of the area under the curve for repeated assessments in a clinical trial [5]. There is some evidence that the value implied by summing the values for each of the health states within a profile does not provide a good approximation of the value of the profile when it is measured directly as a whole [6]. If these assumptions are violated, then this QALY algorithm may give a false impression of the value associated with different health profiles, and hence the relative effectiveness of different interventions might then be misrepresented [7].

Despite recent interest in the issue of whether summing QALY scores derived from constituent health states of a profile gives an accurate value for a profile, there has been relatively little empirical research in this area for use in economic evaluation. Research in the QALY area has focused largely on the sequence effects of the constituent health states. Richardson and colleagues [6] found that profiles, which deteriorate over time, are valued lower than those which vary little over time. Lipscomb found that profiles which improve with time, and finish on a relatively highly valued state, are valued higher than those which deteriorate or show little variation [8]. This suggests that the sequences of states has a systematic effect on the value of a profile. However, a study by Mackiegan and colleagues showed that, where there is very gradual deterioration, sequence has no significant effect on profile valuations [9].

There have been more investigations of preferences over time in the psychological literature. There is evidence, for example, that people prefer profiles that improve over time and end on a good note to profiles with the reverse pattern [10]. It has also been found that preferences for improving sequences are moderated by prior expectations of what is realistic [11]. A more general explanation relevant to the situation of IBS is that profiles may contain gestalt points, which are given particular weight by the valuer. In particular, the final point, the peak point, and the rate of change in measurements over the profile have been found to be significant predictors of the overall rating of the profile [12–14].

The psychological literature has shown that each part of the profile is not of equal value. Despite this research into the effects of duration and sequences on preferences, there has been very little research in the context of health and specifically the QALY assumptions of linearity over time. There has also not been an investigation of these phenomena in conditions such as IBS that offers a particularly interesting example due to the variable nature of the condition.

The main aim of this study was to compare directly elicited values for IBS health profiles with the values implied by the QALY algorithm for the same profiles. Should the study find direct and indirect valuations of the scenarios be consistent, then this implies that none of the violations described above have occurred. However, should they prove to be inconsistent in a statistically and decisionally important way, our analyses will be able to identify the discrepancy but not its determinants.

Methods

The IBS health states and profiles

For this study IBS has been described in terms of two symptoms: whether or not they had experienced adequate relief of abdominal pain and discomfort (P) in the last seven days, and for the number of those days on which they felt a sense of urgency (U). Urgency is described as a dichotomous variable of more than three days and three days or less in order to construct health states. In addition, there was interest in constipation (C) as a side effect of treatment. These three key symptoms form eight possible health states. Of these, six were selected for use in this study to keep the number of valuation exercises to a minimum to avoid cognitive overload. The two excluded states

contained symptoms of constipation. The six states considered in this study are presented in Table 1. Hereafter, the presence of a symptom is indicated by a '+' sign and the absence of the symptom by a '−' sign; thus $P-U-C-$ is the best IBS-related health state and $P+U+C+$ the worst state. It is important to bear in mind that being in the best IBS-related health state does not imply the individual is in full health.

The aggregation of health states to form time profiles of health was less straightforward. Previous work constructing health profiles has tended to deal with scenarios where patients are expected to go through a clear sequence of states. However, IBS does not follow a clear course of progression. Rather, it is characterised by considerable fluctuations in HRQoL over relatively short periods. In addition, symptoms come and go in patterns that vary widely across patients. One way of summarising the patient's health experience using the eight IBS health states is in terms of the proportion of time spent in each state over a set period. A successful treatment may not entirely eliminate the patient's experience of a given state, but may reduce the amount of time spent in that state. The profiles are described in terms of the numbers of weeks in which a patient would be in each IBS state over a 12 week period, which will be repeated for the remainder of the respondent's life. There are four profiles used in this study (Table 2). The interview emphasised that it is not known precisely how the overall length of time in each state is distributed over the 12 week period.

Valuation survey

The valuation exercise. The standard gamble (SG) technique has been used to elicit preferences for the health states and profiles [15]. Basically, the SG asks the respondent to choose between the certainty of an intermediate health state and the uncertainty of a treatment with two possible outcomes, one of which is better than the certain outcome and one of which is worse. The worst state is usually immediate death, though it does not need to be. The health states in this survey are described as lasting for the rest of their life. The objective is to find the probability, p, at which the respondent is indifferent between accepting the intermediate health state and accepting the uncertain treatment. The value for the intermediate health state is then simply given by p.

Table 1. The six IBS health states

Health-state description	Code
You have adequate relief of abdominal pain and discomfort. You do not feel a sense of urgency more than three days per week.	$P-U-C-$
You do not have adequate relief of abdominal pain and discomfort. You do not feel a sense of urgency more than three days per week.	$P+U-C-$
You have adequate relief of abdominal pain and discomfort. You feel a sense of urgency more than three days per week.	$P-U+C-$
You do not have adequate relief of abdominal pain and discomfort. You feel a sense of urgency more than three days per week.	$P+U+C-$
You do not have adequate relief of abdominal pain and discomfort. You do not feel a sense of urgency more than three days per week. You experience constipation.	$P+U-C+$
You do not have adequate relief of abdominal pain and discomfort. You feel a sense of urgency more than three days per week. You experience constipation.	$P+U+C+$
Immediate death	

Table 2. The health profiles

	Number of weeks (out of 12) spent in each state in profile			
Health states	A	B	C	D
$P+U+C-$	2	4	6	2
$P+U-C-$	2	2	2	2
$P-U+C-$	2	2	2	2
$P-U-C-$	6	4	2	6
Constipation	—	—	—	2[a]

[a]These two weeks are distributed over the 12 weeks in ways described in the text.

The version of SG used in the study was originally developed by Jones-Lee and colleagues [16]. This version lists values for the chances of success from 0 in 100 to 100 in 100 and respondents are asked to place a tick against all the probabilities of success at which they are confident they would choose the treatment and a cross against all the values where they would reject the treatment. The space between the ticks and crosses indicates the region of indifference. If this region covers more than one probability value, then a mid-point is taken to be the indifference value. This version of SG has been found to produce more consistent and reliable data than an interview-based variant using props [15].

Many of the IBS health states would be regarded as comparatively mild and so many respondents might be unwilling to accept even a very small chance of dying should the treatment fail. To improve the sensitivity of the SG question it was decided to change the worst treatment outcome in the SG question from death to $P+U+C-$ for valuing two of the states (i.e. $P+U-C-$ and $P-U+C-$) and three of the profiles (i.e. A, B and C). State $P+U+C-$ was in turn valued against the best state and death. This allowed for the two states and three profiles valued against $P+U+C-$ to be 'chained' onto the conventional full health-to-death scale used for calculating QALYs. If the health state value obtained in the first gamble is U_1 and the value of the reference state $(P+U+C-)$ from the second gamble is U_2, the formula for chaining onto the 0–1 scale is simply: $U_3 = U_1 + (1 - U_1)(U_2)$.

The sample. The inclusion criteria for this valuation study were women aged 18 or over who had been diagnosed with IBS based on the Rome criteria. The aim was to recruit a sample of 50 respondents from four ambulatory care centres in UK. This sample size was based on the ability to detect a difference of about 0.1 (on a 0–1 scale) between different health states for an alpha of 5% and power of 80%. There is no real consensus about what difference is to be considered meaningful but a difference of this kind is likely to be important in many decision-making contexts [17].

The interviews. Those eligible for inclusion in the study were interviewed in small groups of 2–6 by a trained interviewer (IT). Whilst some studies have elicited health state valuations using postal questionnaires, there is a great deal of evidence to suggest that this is an unreliable way to gather such data, not least because respondents often interpret questions differently than intended [18]. By bringing respondents together in small groups, it was possible to check that they understood the self-completed questionnaire. Respondents in the group were asked at the end of a detailed explanation given by the interviewer whether they understood the task. They were encouraged throughout to ask for clarification at anytime. The risk of a 'bandwagon' effect was minimal since the questionnaire was self-administered and the interviewer was there to answer questions. In practice, there was little dialogue between respondents. Although this process made it possible for the interviewer to influence the valuations of respondents, it was decided that this risk was worth the gain in respondent understanding. A pilot study was also conducted to ensure that the interview procedure was understood as intended.

Respondents were taken through a questionnaire, which was to be self-completed. The questionnaire began by asking respondents for some background information relating to their age, occupation, age when they completed educa-

tion, the length of time they had experienced IBS for, and a general health question. They were then asked to indicate their current experience with the three IBS symptoms. It was made clear to them all that these questions formed the basis of the health states they would be asked to value. At this stage, respondents had the opportunity to ask questions relating to how the descriptions should be interpreted.

Respondents were then asked to rank the states and profiles in order to familiarise themselves with the descriptions. They were then taken through the first SG question in order to help them understand the exercise, at which point they were given the opportunity to ask any questions. Once they were satisfied with what they were being asked to do, they were asked to complete the remainder of the questionnaire. The questionnaire asked them to value five IBS-related states (the two mildest against a treatment failure state of $P+U+C-$ and the remaining three against death) plus four profiles. The best state formed the upper anchor and so was not valued.

Respondents were also asked to rate their own health using EQ-5D [19,20]. Once everybody in the group had completed the questionnaire, respondents were provided with the opportunity to comment on the exercise.

The analysis of the study data

The first phase of analysis concerns the quality of the data. An important check on the extent to which respondents were able to understand the valuation tasks is to examine the logical consistency of their SG responses. For some pairs of IBS health states, one state can be regarded as logically better than another if it is better on one or more of the symptoms and no worse on any symptom. It is possible for such pairs of states to have the same value, but the logically better state should not be given a lower score. Similarly, for pairs of time profiles, one profile is logically better than another where it contains less time in one symptomatic state and no more time in an other symptomatic state. The number of patients in the sample who did not meet this definition of consistency was calculated.

To calculate the implied profile values using the QALY algorithm, each of the constituent health state values is multiplied by the overall proportion of time spent in that state. This is straightforward for profiles A–C, which do not contain periods of constipation. However, profile D includes two weeks with constipation but it does not specify when this might occur. One solution is to assume that the impact of each symptom is additive, and thus the state of constipation should be given the same value regardless of what other symptoms are present. The periods of constipation were therefore assigned to one week of state $P+U+$ and one week of $P+U-$. However, the impact of constipation on health state values may not be additive.

The statistical analyses included the calculation of the descriptive statistics for the values of IBS states and profiles, and differences between direct and indirect valuations were examined by *t*-test.

Results

In total, 49 respondents took part in this study. The interviews took between 30 min and 1 h to complete, with larger groups taking longer. The mean age of the sample was 47 and the average number of years with IBS was 25. Of the women, 74% were in paid employment. In terms of the IBS symptoms, 34% reported pain and discomfort, 72% felt a sense of urgency for more than three days in the past seven days, and 36% experienced constipation. The mean EQ-5D generic health status score for the sample using Tariff 1 from the York MVH survey of the UK general population was 0.66 [20].

The logical consistency conditions in the health state valuation component of the study are $P+U+C- > P+U+C+$ and $P+U-C+ > P+U+C+$. There are very few violations of the consistency conditions relating to these health states, with rates of 10 and 8%, respectively, and these are comparable with other studies [11]. It is noteworthy that there are more violations of consistency in the direct profile values. The logical consistency conditions are $A > B$, $A > C$, $B > C$ and $A > D$, with 18, 23, 6 and 25% violations, respectively. However, many of the inconsistencies were such that the value of the logically worse state or profile was within 0.05 of the value of the logically better state or profile at the individual level. In addition, the inconsistencies were unrelated to respondent characteristics and no respondent was inconsistent throughout.

Table 3 shows the mean values of the five health states, where one is the best state defined by absence of IBS symptoms (i.e. state $P-U-C-$). Since one respondent had missing data for health state $P+U+C-$, and some values had to be chained to death through this state, it is sometimes only possible to produce results for 48 respondents. It can be seen that the overall values follow a logical ordering and that all mean values are significantly below the value for full health, indicating that the presence of IBS symptoms had a significant effect on quality of life.

Table 4 compares the directly elicited profile values with the implied values associated with those same profiles as calculated using the QALY assumption of linearity over time. The directly elicited profile values were found to be less than the implied value for profiles A and D, and more than the implied value for B and C. For profile A the difference is nearly significant ($p = 0.07$), but nowhere near so for profile B. Differences in profiles C and D were significant at the 5% level.

For the three profiles not containing constipation (i.e. A, B and C) these results show a reversal of the relationship between direct and implied profile valuations, with the direct valuation being less than the indirect valuation for profile A and more for profile C and nearly the same for profile B. This reversal results from the fact that the implied and direct valuations of the profiles follow a different pattern. Direct profile mean and median values are essentially the same across these three profiles (with the value of the health decrements from profile A being −0.002 and 0.000 for profiles B and C, respectively, which were non-significant at the 5% level), whilst the implied values decline

Table 3. Health state values

Health state	*N*	Median	Mean (SD)	Mean difference from 1.0
$P-U+C-$	49	0.997 (0.988–1.000)	0.982 (0.04)	0.018*
$P+U-C-$	49	0.997 (0.990–1.000)	0.986 (0.03)	0.014*
$P+U+C-$	48	0.995 (0.980–1.000)	0.952 (0.10)	0.048*
$P+U-C+$	49	0.995 (0.975–1.000)	0.952 (0.12)	0.049*
$P+U+C+$	49	0.985 (0.965–1.000)	0.933 (0.17)	0.067*

*Difference is significantly different from 1.000 at $p<0.05$.

Table 4. Direct and implied profile values

Profile	*N*	Median (IQR)		Mean (SD)		95% confidence interval around the difference	Paired *t*-test (2-tail significance)
		Direct	Implied	Direct	Implied		*p*
A	48	0.997 (0.992–0.999)	0.998 (0.992–0.999)	0.982 (0.05)	0.987 (0.03)	−0.001 to 0.011	0.076
B	48	0.997 (0.992–0.999)	0.997 (0.987–0.998)	0.980 (0.05)	0.979 (0.045)	−0.002 to 0.004	0.513
C	48	0.997 (0.991–0.999)	0.996 (0.983–0.997)	0.982 (0.04)	0.970 (0.06)	0.005 to 0.019	0.002
D	48	0.985 (0.975–0.995)	0.997 (0.990–0.998)	0.937 (0.16)	0.982 (0.03)	−0.083 to −0.007	0.021

in a linear fashion as would be expected given the QT QALY assumption (with the value of the health decrements from profile A of profiles B and C being 0.008 and 0.017 ($p < 0.05$)).

Given the comparatively high levels of respondent inconsistency, the same analysis of mean values was undertaken excluding the 15 respondents who were inconsistent in any of the six ways identified earlier. This did not alter the pattern of the relationship between the direct and implied valuations.

Profile D includes the additional symptom of constipation, and therefore does not enter into the above comparison. It has the same amount of time in pain and urgency as profile A, but has two unspecified weeks associated with constipation, which run along side the other symptoms and could occur at any point during the 12 weeks. The mean profile value for profile D is significantly less than the mean value of profile A. The direct valuation was also found to be significantly lower than the implied valuation.

Discussion

This paper has an important contribution to make to the literature on the valuation of health for use in economic evaluation. It has been claimed in the past that the QT QALY algorithm does not adequately represent people's preference over time profiles of health [7]. The study reported in this paper would appear at first sight to provide evidence of a violation of the linearity assumption of the QALY model.

Despite the higher levels of inconsistency and dispersion associated with direct valuations it is hard to dismiss this finding as simply an artefact. The result was robust to the exclusion of inconsistent respondents and was found to be common across most respondents at the individual level. The general health status score of the sample indicated states of health significantly below that expected in people of their age [21]. Furthermore, many of them were experiencing the symptoms they were being asked to value. However, there may be a concern that the sample of IBS sufferers were involved in an ongoing clinical trial and may not be perfectly representative of IBS sufferers in the community at large. This was mainly a methodological study and hence the precise sample of respondents used is not important.

The SG task did not have death as one of the reference states for five out of the nine valuations. This was done for two reasons. The first was that the valuation task needed to be realistic for the respondents to take them seriously. There was also a concern that respondents might become alarmed at a reference to death in every SG question. Therefore, for the milder states and profiles, the reference states of $P-U-C-$ and $P+U+C-$ have been used, where the latter was worse than the states being valued. The second reason was to increase the sensitivity of the scale. However, this paper is addressing issues around the derivation of QALYs, where zero represents states regarded as equivalent to death and one states in full health, and it has been necessary to chain the values using the valuation of $P+U+C-$. This can be criticised due to inconsistency between chained values and direct valuations [22,23]. However, the analysis has been repeated for profiles A, B and C using unchained values and the same relationship has been found between direct and implied profile values.

It is important for possible users of the results presented in this paper to appreciate that the symptom-free state $P-U-C-$ does not necessarily correspond to full health. Whilst we were able to chain the health state and profile valuations to death, the upper end of the scale could not be anchored to the end point of full health. This does not matter for the methodological comparison presented in this paper, but the results could not be used in cost-utility analysis without further adjustment [24].

The fact that $P-U-C-$ is not full health may also partly explain the discrepancy between mean IBS health state values of 0.933–0.986 and the respondents' own EQ-5D scores of 0.66. The IBS states take no account of other aspects of the condition or co-morbidity. Furthermore, the SG valuations of the IBS states will be subject to risk aversion. Finally, the EQ-5D scores have been obtained from a general population sample and these have been shown to be lower than those of patients in many conditions [25].

This paper offers an important insight into the way people value conditions such as IBS, where the condition does not follow a clear progression. Previous work in the economics literature on valuing time profiles of health has tended to focus on simple sequences of states. IBS is characterised by two distinct features, which have not been examined in previous work on valuing profiles.

The first is that IBS is subject to considerable and unpredictable fluctuations in its symptoms and hence HRQoL varies over relatively short periods. Yet at the same time it is a condition that can last for many years. We developed a method for summarising the patient's health experience using four IBS health states in terms of the proportion of time spent in each state over a set period (12 weeks), which will be repeated for the remainder of the respondent's life. The time spent in each health state is presented as being scattered in an unpredictable way over each period of 12 weeks. This method allows the outcomes of treatment to be compared in terms of the average amount of time spent in each state. This study showed how such profiles can be valued directly using SG. IBS is not the only condition with these features and the results presented in this paper have implications for other medical conditions.

This study adds to the currently limited amount of empirical work in the economic literature addressing an important debate between the use of health state values via QALYs and the alternative of the HYE or profile valuation. The results provide evidence of the incongruence between the QALY assumption and the way people directly value time profiles of health. Whilst the inconsistency was significant in three out of the four profiles examined, the differences were rather small at 0.005 (A), 0.001 (B), 0.012 (C) and 0.045 (D). It could be argued that such small differences would have little decisional impact. However, these differences need to be measured against a very narrow range of values for these profiles, which range from 0.937 to 0.987. For a patient choosing between treatments, these differences may still prove to be important. The economic importance of these differences really depends on context and while the differences may be small in absolute terms, they may nonetheless prove to be important where the cost differences are also small.

However, this study was not designed to explain the differences found. The QALY model used in this study to value the profiles (i.e. the 'implied' profile values) essentially assumes, following on from the distinctions made by Gafni and Torrance, no sequence effects, no quantity effects and a zero time preference [26]. The results found in this study could be partly explained by anyone of these. The sequence of states in the profile was not specified for the direct valuation. This was left to the respondents to think of possible sequences. The sequence of states would also have implications for time preference. However, given that we know neither the sequence imagined by the respondents nor their time preference rate, we do not know the likely direction or size of these effects.

It could be a result of a quantity effect whereby individual valuations of states is not proportional to the time spent in them. There could be some kind of threshold effect causing a non-linear pattern in direct valuations. This hypothesis can only be tested through further research valuing IBS profiles with 8, 10 and 12 weeks in state $P + U + C-$.

There are other explanations found outside of the health economics literature. One possible explanation might be found in the work of Ariely and Zauberman who addressed issues around valuing a whole profile versus segments [27]. They showed that the overall rating of an entire profile of non-health-related events such as annoying sounds is based on the gestalt characteristics; but when the profile is broken into segments which are rated separately, the mean value of each segment becomes more important in the overall valuation of the profile. The work by Ariely and Zauberman could go some way towards explaining the difference between implied and direct methods of valuing IBS health profiles. The implied values are obtained by segmenting the profiles into their constituent health states. However, the direct valuations are based on gestalt characteristics, which for the IBS profile could simply be the worst health state, rather than the proportion of time they could expect to be in the state. For valuations by the implied method a good segment would be valued highly, and a bad segmented valued low. This would suggest that the implied method would value a profile that tended to be good more highly and a profile with a tendency to be bad more lowly. This may account to some extent for the steps in the valuation of profiles A–C found in this study.

It has been argued by some commentators that the direct valuation of profiles is superior to the QALY algorithm since it does not assume people's preferences are linear over time. This position is based on conventional choice theory and takes no account of the cognitive difficulty associated with the valuation of profile compared to health states. The greater difficulty associated with valuing profiles is reflected in the higher inconsistency rates. The psychology literature suggests that respondents in this circumstance may use simplify-

ing heuristic or gestalt devices to overcome the cognitive load, such as concentrating on the state, rather than considering the time spent in the state. For the purpose of valuing benefits of health care, it will be necessary to address these problems with the direct valuation of profile. Establishing which method offers a better way of valuing health profiles ultimately requires some test of external validity, but this requires further research.

Acknowledgements

We are grateful to Glaxo-Smith-Kline for funding this study. We would like to thank the reviewers for their constructive comments. The usual disclaimers apply.

References

1. Ministry of Health (Ontario). *Ontario Guidelines for the Economic Evaluation of Pharmaceutical Products*. Ministry of Health: Toronto, 1994.
2. Commonwealth Department of Health, Housing and Community Service. *Guidelines for the Pharmaceuticals Industry on the Submission to the Pharmaceutical Benefits Advisory Committee*. Australian Government Publishing Service: Canberra, 1992.
3. Department of Health and the Association of the Pharmaceutical Industry. *Guidelines for the Economic Evaluation of Pharmaceuticals* [*Press release*]. Department of Health: London, 1994.
4. Bleichrodt H, Wakker P, Johannesson M. Characterizing QALYs by risk neutrality. *J Risk Uncertainty* 1997; **15**: 107–114.
5. Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *Br Med J* 1990; **300**: 230–235.
6. Richardson J, Hall J, Salkfeld G. The measurement of utility in multiphase health states. *Int J Technol Assess Health Care* 1996; **12**: 151–162.
7. Mehrez A, Gafni A. Quality-adjusted life years, utility theory, and healthy-years equivalents. *Med Decis Making* 1989; **9**: 142–149; **10**(2): 148–149.
8. Lipscomb J. The preference for health in cost-effectiveness analysis. *Med Care* 1989; **27**: S233–S253.
9. Mackeigan LD, O'Brien BJ, Oh PI. Holistic versus composite preferences for lifetime treatment sequences for type 2 diabetes. *Med Decis Making* 1999; **19**: 113–121.
10. Varey C, Kahneman D. Experiences extended across time: evaluation of moments and episodes. *J Behav Decis Making* 1992; **5**: 169–185.
11. Chapman GB. Preferences for improving and declining sequences of health outcomes. *J Behav Decis Making* 2000; **13**: 203–218.
12. Ariely D, Carmon Z. Gestalt characteristics of experiences: the defining features of summarized events. *J Behav Decis Making* 2000; **13**: 191–201.
13. Kahneman D, Fredrickson BL, Schreiber CA, Redelmeier DA. When more pain is preferred to less: adding a better end. *Psychol Sci* 1993; **4**: 401–405.
14. Redelmeier DA, Kahneman D. On the discrepancy between experiences and memories: patients' real-time and retrospective evaluations of the pain of colonoscopy. *Working Paper*, 1993.
15. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. *J Health Econ* 1996; **15**: 209–231.
16. Jones-Lee M, Loomes G, O'Reilly D, Phillips P. The value of preventing non-fatal road injuries: findings of a willingness-to-pay national sample survey. Transport Research Laboratory, 1993.
17. O'Brien BJ, Drummond MF. Statistical versus quantitative significance in the socioeconomic evaluation of medicines. *PharmacoEconomics* 1994; **5**: 389–398.
18. Slovic P. The construction of preferences. *Am Psychol* 1995; **50**(5): 364–371.
19. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; **37**: 53–72.
20. Dolan P. Modelling valuations for EuroQol health states. *Med Care* 1997; **35**: 351–363.
21. Kind P. The EuroQol instrument: an index of health-related quality of life. In *Quality of Life and PharmacoEconomics in Clinical Trials* (2nd edn), Spilker B (ed.). Lippincott-Rivera: Philadelphia, PA, 1996; 191–201.
22. Rutten-Van-Mölken M. Costs and effects of pharmacotherapy in asthma and COPD. *PhD Thesis*, Universitaire Press Maastricht Datawyse Maastricht, 1994.
23. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. Describing health states: methodological issues in obtaining values for health states. *Med Care* 1984; **22**: 543–552.
24. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes* (2nd edn). Oxford Medical Publications: Oxford, 1997.
25. Brazier JE, Akehurst R, Brennan A *et al*. Should patients share a greater role in valuing health states? *Health Economics and Decision Science Discussion Paper Series no. 04/3*, ScHARR, University of Sheffield, 2004.
26. Gafni A, Torrance GW. Risk attitude and time preference in health. *Manage Sci* 1984; **30**: 440–451.
27. Ariely D, Zauberman G. On the making of an experience: the effects of breaking and combining experiences on their overall evaluation. *J Behav Decis Making* 2000; **13**: 219–232.

