CrossMark

CURRENT OPINION

Improving Cross-Sector Comparisons: Going Beyond the Health-Related QALY

John Brazier[1] · Aki Tsuchiya[1]

Abstract The quality-adjusted life-year (QALY) has become a widely used measure of health outcomes for use in informing decision making in health technology assessment. However, there is growing recognition of outcomes beyond health within the health sector and in related sectors such as social care and public health. This paper presents the advantages and disadvantages of ten possible approaches covering extending the health-related QALY and using well-being and monetary-based methods, in order to address the problem of using multiple outcome measures to inform resource allocation within and between sectors.

Key Points for Decision Makers

Quality-adjusted life-years (QALYs) are traditionally health related and so do not cover all the things that may matter to decision makers in health and related sectors.

There are numerous alternatives from extending the existing health measures to using a broader notion of well-being or monetary-based approaches.

Ultimately, there are important value judgements to be made about what matters in public policy.

✉ John Brazier
   j.e.brazier@sheffield.ac.uk

[1] School of Health and Related Research (ScHARR),
    University of Sheffield, 30 Regent Street, Sheffield S1 4DA,
    UK

1 Background

The quality-adjusted life-year (QALY) has become a widely used measure of outcome for use in informing decision making in health technology assessment. QALYs are commonly estimated using preference-based measures such as the EQ-5D, SF-6D and HUI3 [1–3]. The most widely used instrument for estimating the quality of life component of the QALYs has been the EQ-5D [4]. EQ-5D is designed to measure an individual's generic health status (or health-related quality of life) across five dimensions: personal functioning (with mobility and self-care); activities (usual activities); pain or discomfort; and anxiety or depression, each with three levels [5]. It has an accompanying preference-based value set obtained from the general public using a variant of time trade-off (TTO) [1]. There is currently research into the valuation of a new 5-level version [6]. EQ-5D is the preferred measure for use in assessing the cost effectiveness of health technologies for NICE [7], and it is used in the NHS Executive's PROM (Patient Reported Outcome Measures) programme.

   However, there is recognition of outcomes beyond health and that measures such as EQ-5D are not adequate in related sectors such as social care and public health. Even in health care, for some conditions EQ-5D does not capture all the things that matter to patients [8]. In social care there has been the development of the Adult Social Care Outcome Tool (ASCOT) for routine use in social services [9]. In public health, there is no single measure but there are a number of broader measures that could be used, including the preference-weighted ICECAP capability index [10, 11], measures of well-being such as the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) [12] and the ONS-4 [13]. Most do not have any preference

△ Adis

weighting; the only one that does (ICECAP) is not anchored at 0 for being dead, which would allow them to be used to estimate the quality adjustment component of the QALY. The use of multiple outcome measures presents decision makers (such as NICE) with the problem of how to use such measures to make comparisons across sectors or how to combine them to provide an overall measure of benefit whilst avoiding double counting.

2 Aim

The UK Medical Research Council (MRC) sought to examine alternatives to the health-related QALY to address the problem of using multiple outcome measures to inform resource allocation within and between sectors. To this end it commissioned this review of the alternatives in order to inform its recent highlight notice to encourage research applications in the area [14]. It does not claim to be a full systematic review, nor has it identified all possible alternatives. This review aims to provide an overview of a representative list of ten approaches, and presents the advantages and disadvantages of each and identifies research required to develop them further.

This review does not go in to details of the methods of valuation, which include TTO, standard gamble, visual analogue scale, discrete choice experiments (DCEs), and person trade-off. The DALY (disability-adjusted life-year) is not reviewed, although its 2010 version now uses general public preferences as the basis of disability weights [15]. This is because its aim is to quantify the burden of disease, or loss of health, as opposed to loss of welfare/well-being [16]. We exclude approaches based on the valuation of whole time profiles of health, such as healthy year equivalents [17], since the focus of this paper is on the use of standardised measures (e.g. EQ-5D, ASCOT and WEMWBS). Finally, this review does not address in any detail the implications for economic evaluation studies.

3 The Main Approaches

The ten main approaches reviewed in this paper are classified under three headings: those looking to extend the existing health-related QALYs, those using well-being and those using money to value outcomes.

Extending the QALY beyond health:

A. Statistical mapping to EQ-5D.
B. Bolting on to EQ-5D.
C. Valuing on a common scale using preferences.

    Using well-being to value outcomes:

D. Valuing by association with well-being.
E. Developing a well-being-adjusted life-year (WELBY).
F. Direct valuation of own health or well-being states.

    Using money to value outcomes:

G. Public sector implied willingness to pay (WTP).
H. Contingent valuation using WTP (welfarist).
I. Societal WTP (non-welfarist).
J. Monetarise health and other outcomes using experience.

3.1 Extending the QALY Beyond Health

A. Statistical mapping to EQ-5D.

Most statistical mapping is from a non-preference-based and/or condition-specific measure to a preference-based generic measure such as the EQ-5D. Mapping is one option recommended by NICE [7] to estimate EQ-5D utility data when EQ-5D data are unavailable in the study dataset. Estimating a mapping function between EQ-5D and another instrument (e.g. ASCOT) would require both to be collected together in one or more datasets. A mapping function could be estimated by regression which would enable any ASCOT state to be linked to an estimated value for the EQ-5D. There are a variety of specifications that can be fitted to the data and different statistical techniques for dealing with the distributions of the variables involved (see examples from health care) [18–20]. This would mean that if an evaluation of a social care intervention collected only the ASCOT measure then these data could be used to estimate EQ-5D values.

However, the mapping function relies on statistical association and this is unlikely to be strong given the low conceptual overlap between ASCOT and EQ-5D. EQ-5D is about the key five aspects of a person's health, whereas ASCOT is concerned with the way a person's health—combined with their socio-economic status, home circumstances (including availability of informal care), and the social care services they receive—impacts their overall quality of life defined in terms of the extent to which their needs and wants are being met. For example, the same (poor) health in EQ-5D can impact a person's ASCOT score in different ways depending on the availability and quality of informal and formal care provision. At the same time, the provision of good social care may result in different levels of EQ-5D health achieving the same ASCOT score. The descriptive systems are simply not measuring the same thing. There are similar problems in trying to map from ICECAP or WEMWBS to EQ-5D.

In these circumstances the EQ-5D would be unable to reflect many of the outcomes captured by the other measure

and so direct statistical mapping would not be an appropriate solution.

B. Bolting on dimensions to EQ-5D.

Another alternative approach when EQ-5D does not cover the relevant outcomes is to expand the measure by including 'bolt-on' dimensions to cover the dimensions missing from the measure. Bolt-on dimensions to EQ-5D have been developed for cognition [21], sleep [22], energy, vision and hearing [8]. Recently, there was a study looking at the addition of satisfaction [23]. The wording of these dimensions tends to conform to the format of the EQ-5D dimensions: no problem, some problem and severe problem. To be useful, these bolt-ons need to be incorporated into the EQ-5D preference-based value set so they can be compared with the impact of the other dimensions and between bolt-ons. There is evidence that the addition of a dimension has consequences for the value of the other dimensions; for example, a vision bolt-on was found to reduce the coefficient on usual activity [8], implying a need for the re-valuation of the EQ-5D value sets with the bolt-on (see item D below for 'generalised' TTO).

This solution has a more fundamental limitation since the problem is often not simply the absence of one or two dimensions. The use of bolt-ons may have potential within health care where there are just one or two missing dimensions, but there is little conceptual similarity between measures such as EQ-5D, ASCOT, ICECAP and WEMWBS. Furthermore, there is a limit to the number of dimensions that a descriptive system can have for it to be amenable to valuation tasks such as TTO or DCE. Furthermore, there is a risk of double counting since the dimensions in one measure may be captured to some extent by dimensions in the other measure; for example, ability to meet personal care needs may be partly reflected in the self-care dimension of EQ-5D. There would seem little value in pursuing research into this approach for tackling cross-sectorial issues.

C. Valuing on a common scale using preferences.

QALYs are based on the elicitation of the preferences of the population for living in different health states. The EQ-5D was valued using TTO, where a respondent is asked to compare a life of 10 years in a given ill health state with a shorter period in full health. The period in full health is varied until the individual is indifferent between health state z for 10 years and full health for x years, at which point the value or the quality adjustment weight of state z is derived as $x/10$ [24]. This means all EQ-5D states are valued on a scale with zero for dead and one for full health. From a theoretical perspective, choice-based methods such as TTO imply that the quality adjustment of the QALY is equivalent to an overall well-being adjustment—everything of value to an individual will be incorporated into it. In a TTO exercise, if the individual is indifferent between health state z for 10 years and full health for 6, each year in this health state is valued at 0.6. The difference between state z and 'full health' is being valued in terms of everything which is important about being alive, as it is not just the 4 years of health which is traded off, but 4 years of life.

Thus, TTO can be argued to capture more than health through the impact of health on overall quality of life, though this is limited by the accuracy with which a respondent is able to imagine these broader impacts (see discussion under approach F on direct valuation). Choice-based techniques like TTO could provide a way to compare measures like EQ-5D and ASCOT, which have been valued using this method. However, TTO tasks used to value these instruments differed in a crucial way—the upper anchor (viz. the better state in shorter duration) tends to be instrument specific: for EQ-5D it was EQ-5D state 11111 (no health problems) and for ASCOT it was ASCOT state 11111111 (meeting social-care-related needs and wants). These upper anchors are not the same, which may result in important differences in the scales.

What is required is a common yardstick. Exploratory research funded by MRC examined the use of a generic Visual Analogies Scale (VAS) (best imaginable to worst imaginable life) and ranking methods to value a number of measures including EQ-5D, an earlier version of ASCOT, ICECAP, and an asthma-specific measure [25, 26]. This enabled the estimation of exchange rates between these measures. This approach could be extended to a choice-based valuation technique such as TTO, where the upper anchor is not instrument specific but described in more general terms such as 'best imaginable state'. In a more explicit way, respondents are being asked to value states defined by EQ-5D and ASCOT in terms of how many years of best imaginable life they would be willing to sacrifice. Once a sample of states from the two instruments have been valued in this way it would be possible to map between them using the common scale.

However, concerns still remain. For example, when valuing EQ-5D the respondent's attention is focused on the particular aspects of health and they are typically not encouraged to think more broadly about their life. It is not clear what they imagine will happen to other aspects of their life like job, income, relationships, well-being and so forth. The nature and extent of this problem could be investigated using mixed methods, including in-depth interviews into what respondents say they are taking into account in the task and empirical work into the impact of these factors in the valuation of states.

3.2 Using Well-Being

The approaches described so far in this document are reliant on the use of preference, elicited using techniques such as TTO where respondents are being asked to imagine health or social care states. These approaches assume that individuals are able to predict the likely impact of the health state being described on their future lives, but this has been shown not to be the case in health and other contexts [27]. General population respondents to health valuation surveys imagining health states, for example, usually do not take into account the extent of any adaptation they may make over time [28]. So their preferences will provide a poor indicator of the actual impact on their well-being. This is one of the reasons why some economists have advocated the use of more direct measurements of well-being in those experiencing the health states through measures of subjective well-being [27].

This raises the issue of what is well-being. Well-being, broadly conceived, is how well an individual's life is going. Subjective well-being (SWB) has been described under three headings: hedonism (well-being increases when an individual experiences more pleasure and/or less pain), flourishing theories (well-being increases when an individual fulfils their nature as a human being, or 'flourishes') and life evaluation or life satisfaction (well-being increases when an individual positively assesses her life) [29]. Traditionally, there are also objective list accounts of well-being including items such as literacy, accommodation and ability to see [30].

The well-being literature has seen the rise of the use of Sen's notion of a capability set. Capability sets are made up of those things you can do or be [31]. He advocated the use of capability sets in response to concerns about an over-reliance on outcomes and utilities in economics. Sen argued that society is interested with what you can do or be, rather than what you actually choose (or happen) to do or be. This contrasts with conventional consequentialist measures like EQ-5D. Although Sen remains reluctant to set out a definitive list of capabilities, there have been several attempts [10, 32, 33]; the problem is that it is doubtful whether capability sets can be measured using questionnaires [34]. An attempt to measure capabilities in health care is the ICECAP [10], which tries to achieve this by asking respondents whether they 'can have…' or 'are able to…'. The content of having achieved as much as they would like is similar in content to measures of psychological well-being.

There are a number of tools available to measure SWB, including simple self-reported items on happiness and life satisfaction, and multi-item measures of psychological well-being such as WEMWBS [12]. In addition, some economists have used TTO or other techniques directly in people suffering ill-health in order to get their preference-based value. One of the issues in applying well-being approaches has been deciding which measure to use, but here we focus on describing the different approaches to using well-being.

D.   Valuing by association with well-being.

Some economists have advocated the measurement of cardinal utility in terms of subjective well-being, such as affect (e.g. pleasure and pain) or life satisfaction [27]. These well-being scales can be used as dependent variables for estimating the impact of measures like EQ-5D or ASCOT on well-being. These studies have used regression techniques to estimate weights for EQ-5D and SF-6D, including self-reported happiness and satisfaction items [35, 36].

The use of well-being measures is currently limited by the fact that they are often unscored single items or, where there are multiple items, they are simply summed together or valued using the output of psychometric techniques like Rasch [37]. This makes interpretation of the scores problematic since they do not provide a cardinal measure. In the future, it may be possible to generate a multi-item instrument which generates a cardinal score that has something close to interval properties. This would provide a basis for comparison across measures on a common scale.

Another limitation with this approach is that the well-being scales are not anchored on the zero to one scale which is required to calculate QALYs. This limits the application in health (and social care) where mortality is often a key outcome. Perhaps more fundamentally it assumes that the association between health and well-being represents causality. This limitation could be addressed by more sophisticated modelling of longitudinal data containing the measures of interest. However, it is unlikely that such data sets exist for measures outside of health. It also relies on an acceptance of the validity of well-being measures for making inter-personal comparison. Further research is required to estimate and develop a well-being scale that is cardinal with interval properties on a scale where zero is anchored on dead, and to examine longitudinal data sets. This would require primary research into the development of a WELBY (as described in the next section), as well as analyses of secondary data sources.

E.   Developing a well-being-adjusted life-year (WELBY).

A WELBY is the same as a QALY measure like EQ-5D, except the descriptive system is concerned with well-being rather than just health-related quality of life. A multi-dimensional well-being classification system like EQ-5D could be formed from measures such as the ICECAP, ONS-4 or WEMWBS and valued using a generic TTO or other preference elicitation techniques.

A WELBY could be used to measure benefits across sectors, permit comparisons of efficiency across sectors, and be used within the existing framework of economic evaluation. It would allow comparisons of the incremental cost per QALY across sectors and could provide a new and radical way of asking questions of the current resource allocation levels between sectors. The disadvantage is that these general subjective well-being measures are less specific and have been shown to have lower levels of sensitivity than more sector-specific measures such as EQ-5D and ASCOT [38]. A more fundamental concern is that, in order to value a WELBY, preferences are being used and these have been criticised for being poor indicators of the way health impacts on a person's life. However, in the context of a WELBY, it is preferences over well-being rather than health and so should be less prone to this problem.

Research should be undertaken to develop a WELBY to be used across the public sector in order to examine the extent of the implications of these limitations. It would be easier to use an existing well-being measure rather than develop an entirely new one, through statistical mapping between a new WELBY scale and existing measures of SWB.

F. Direct valuation of own health or well-being states.

This approach asks patients and other individuals who are experiencing any given state to value it. It is an approach that avoids the need for a description of health, such as the EQ-5D, or for well-being, such as the WEMWBS. It could be done using a well-being scale, but as already mentioned this has the problem of not being commensurate with survival which is a major limitation in health care. Preference elicitation techniques such as TTO can be used in order to anchor responses onto the full well-being—dead scale used to calculate QALYs. This and other techniques have been used in healthcare patient populations, where the upper anchor tends to be full health [28], but it has not been used in other sectors.

This approach gets to the heart of an important normative issue as to whether resources should be allocated on the basis of some aggregated societal valuation, as is the current method, or some aggregation of values from actual users of the services. The argument for users such as patients is that they know the impact of their own state on their lives better than someone trying to imagine it using the highly stylised descriptions of measures like EQ-5D or well-being scales like WEMWBS. For example, people tend to under-estimate the extent to which they can adapt to physical health states in the longer term, and so provide significantly lower values [39].

Another concern is that well-being is prone to adaptation, resulting in low expectations and even denial effects.

Direct value elicitation has also been criticised on the grounds that it involves imagining the best state and for someone with long-term conditions this is quite hypothetical [39]. There are also practical problems with obtaining values from a representative sample of users, including those who are in poorer states who may be unwilling to perform such a task and indeed it may be unethical to ask individuals in terminal states the life and death questions involved in these elicitation techniques.

This approach has been used for many years and there are some reviews of the results compared with general population samples [38]. However, it has not been used extensively and systematically across patient and other service user populations. Further research could examine the use of this approach more extensively and what it says about the way people value their own state.

3.3 Using Money to Value Outcomes

Traditionally, economists have sought to undertake cost–benefit analyses where all the costs and benefits are valued in monetary terms. This enables the use of the decision rule that an intervention should go ahead where there is a net benefit in monetary terms [3]. This requires all benefits, including any impact of health and well-being, to be valued in monetary terms. This would enable the benefits in one sector to be compared to another, though care would need to be taken to ensure that double counting is avoided. There are a variety of techniques for doing this and here we outline a few approaches that may be considered in the context of this review.

G. Public sector implied willingness to pay (WTP).

This approach utilises the fact of each government department having its own annual budget constraint to achieve their objectives. Under a constrained budget, decision makers are allocating resources to competing demands and these decisions imply relative values to different outcomes. The value of different outcomes is revealed by the decisions being taken by policy makers, whether or not these are optimal. The best example of this is the cost per QALY threshold range used by NICE to inform their decisions to recommend the funding of new health technologies of £20,000–30,000 [7]. Recent empirical evidence indicates that actual decisions made by the NHS suggest the value may be significantly less than this [40]. The threshold is interpreted to represent the amount that relevant decision makers are willing to pay for a given outcome. The research undertaken at the University of York examined natural variation in expenditure across the NHS and natural variation in mortality outcomes to estimate these implied values. In principle, such threshold values exist (if only implicitly) across other sectors, though

there is little available evidence at present to estimate the amounts. Once such values are available across sectors, this approach can provide values for different outcomes across sectors that can be used within a net benefit framework [41]. When one sector (e.g. NHS) generates benefits relevant to another sector (e.g. social services), the beneficiary can compensate the sector producing the benefit. In reality, it may not be possible to undertake actual compensation, in which case a compensation test can operate: if one sector could hypothetically compensate another sector then the intervention/programme should go ahead.

This approach presents some major empirical challenges, since other sectors do not have a threshold like NICE. However, a simplified version of this framework was developed by the Department of Health for Value Based Pricing using a range of shadow prices for outcomes in different sectors [42]. A criticism of using implied values is that they may not reflect the values of society. Furthermore, outcomes across sectors are not unrelated and so there is a risk of double counting. Taking the example of EQ-5D and ASCOT, whilst they differ in many ways, there is a significant degree of overlap in the dimensions they cover and so they cannot simply be combined to generate a total estimate of value.

The potential research required to take this approach forward includes the explicit estimation of thresholds for different outcomes across sectors (e.g. in social care) and an investigation into the optimality of the implied values.

H. Welfarist WTP.

Public sector implied WTP above relies on the assumptions that (a) actual resource allocation in the public sector is efficient and (b) it reflects what the public want. A more traditional approach in economics would be to value non-market goods and services such as environmental changes via 'compensating variations'. This is the maximum amount of money that, following a good change, an individual can pay and still maintain the level of welfare they had prior to the change; thus the term 'willingness to pay' (WTP). Typically, they will be elicited through questionnaire surveys tapping into respondents' decision utility and preferences, using hypothetical scenarios to value [43]. Compensating variation is a cardinal measure of change in individual well-being, as assessed by the individual themselves, and while the underlying individual utility functions are not directly observed, the size of compensating variations appear to be interpersonally comparable. This has been challenged, because it seems intuitively 'wrong' to say a compensating variation of, for example, £100 means the same thing to a rich person and to a poor person. Most welfare economists agree in principle that to be interpersonally comparable (viz. to be of policy

relevance), compensating variation should be corrected for income inequality by applying 'distributional weights' [44].

There are a number of different ways WTP could be used to value cross-sector outcomes. The use of WTP in health care tended to be mainly about the valuation of specific interventions where the respondent is provided with information about benefits including health, convenience and the processes of care [45–47]. There have been some applications to value a QALY through people's WTP to avoid some duration or risk of a health state using EQ-5D [48]. However, the WTP method has not been used to model monetary values for entire descriptive systems to date [49]. Experience with NICE submissions is that the use of such vignettes describing health and other benefits is open to manipulation and carries less weight than one based on patient-reported outcomes [7]. It has not been applied on a large scale like QALYs and it would require a significant amount of work to operationalise.

More generally, there is evidence from the WTP literature, mostly in environmental economics, regarding insensitivity to scope and scale (i.e. things that WTP should respond to) and sensitivity to framing (i.e. things that WTP should not respond to) [50–52]. Furthermore, there are concerns with using WTP to value healthcare services in a system such as the UK NHS, where we do not pay for health care out of pocket. It should also be noted that there has been limited support for the use of WTP in the health sector as a measure of change in individual well-being.

I. Non-welfarist WTP.

Compensating variation above is a concept based on welfarist and individualist welfare economics. However, monetary measures of health and well-being need not be welfarist. Members of the public can be asked how much money government should provide towards different policies, where outcomes might be described in terms of any of the descriptive systems described above. The metric will be directly comparable across sectors. Crucially, the payment is not out of individual pockets, and so distributional concerns are not an issue.

Some sporadic studies exist (e.g. [53]), but not as a coherent body of literature, and this approach is the most 'blue-sky' of all approaches discussed in this paper. In terms of economic evaluation, a non-welfarist variant of the cost benefit analysis will be called for. It would not be appropriate to use net benefit rules in a budget-constrained public sector and so there will emerge cost per WTP thresholds across each sector to inform investment and disinvestment decisions as being used by NICE. However, this information would provide a basis for making comparisons between sectors.

J. Monetarise health and well-being using experience.

Another way to avoid the limitations of a lack of anchoring of well-being and to improve interpersonal comparability would be to estimate the relationship between health and other outcomes in terms of income (i.e. to monetise them) using a well-being equation. This approach converts outcomes into money, but uses the relative impact of the outcome measure compared with income on well-being to provide the calibration. The method models the determinants of subjective well-being in terms of health and other outcomes, and estimates exchange rate, or marginal rate of substitution, between income and health or social care with well-being as the dependent variable. This approach has been used in the context of health [54], but not social care or other outcomes. Furthermore, monetary equivalent values change depending on the well-being measure used [55]. Such modelling needs to take into account the complexity of the relationship since it is not uni-directional. This approach need to be explored further.

4 Conclusion

There are ten approaches to address the problem raised of how to compare outcomes across sectors and avoid double counting. The approaches start from minor adjustments to current methods and progress to options that depart in more radical ways from the health-based QALY. These ten approaches can be divided into three broad sets. The first are those that would represent the least deviation from the current practice of many agencies around the world of using health-related QALYs in health care, by proposing to either map other measures onto preference-based health measures such as EQ-5D, to bolt on dimensions to the EQ-5D as required, or to value all measures using a common generic version of TTO. The second set of approaches uses well-being in different ways to value outcomes across sectors: by association with well-being measures, by direct valuation of own health or well-being using TTO (or some other preference-elicitation technique), or by developing a WELBY. The final set of four approaches all use money as the metric: those implied by decisions in the public sector, contingent valuation using an individual's stated WTP, WTP from a societal perspective and monetising in terms of the impact on well-being. Whilst these alternatives are not exhaustive, they represent the range of alternatives from the least to the most radical. There are other variants of the methods and there are numerous detailed technical issues about how they are to be implemented.

These approaches are not entirely mutually exclusive. An extended health or well-being approach, for example, could be used to describe the outcomes, but they can be valued using a QALY type model (i.e. on the zero to one scale) or monetised using various forms of WTP. Another example is that any new well-being measure could be mapped onto existing measures.

Any choice between these approaches involves important political decisions about what counts in measuring the benefits of interventions. An important example would be the choice of whether to use a pure subjective well-being measure such as happiness or life satisfaction to describe the benefits, compared with more sector-specific outcomes like EQ-5D or ASCOT. It may be that in the end policy makers opt for a compromise involving subjective well-being for the cross-sector comparisons but continuing to use the sector-specific outcomes for within-sector decisions since they are more sensitive. Furthermore, sector-specific outcomes could be valued using subjective well-being (e.g. approaches D or J) or money (e.g. H, I or J).

Compliance with Ethical Standards

Author contributions JB and AT both contributed to the conception and planning; and drafting and revision of the paper. JB is the guarantor for the overall content.

References

1. Dolan P. Modelling valuations for EuroQol health states. Med Care. 1997;35(11):1095–108.
2. Brazier J, Roberts J, Deverill M. The estimation a preference-based single index measure for health from the SF-36. J Health Econ. 2002;21(2):271–92.
3. Feeny D, Furlong W, Torrance GW, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care. 2002;40(2):113–28.
4. Wisløff T, et al. Estimating QALY gains in applied studies: a review of cost utility analyses published in 2010. PharmacoEconomics. 2014;32(4):367–75.

5. Brooks R, The EuroQol Group. EuroQol: the current state of play. Health Policy. 1996;37:53–72.

6. Herdman M, Gudex C, Lloyd A. Development and preliminary testing of the new five level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727–36.

7. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. 2013. http://www.nice.org.uk/article/pmg9/chapter/Foreword. Accessed 23 Aug 2015.

8. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, Rowen D, Tosh J, Tsuchiya A, Evans P, Keetharuth AD, Brazier J. Use of generic and condition specific measures of Health Related Quality of Life in NICE decision making. Health Technol Assess. 2014;18(9):1–224.

9. Netten A, Burge P, Malley J, Potoglou D, Towers A, Brazier J, Flynn T, Forder J, Wall B. Outcomes of social care for adults: developing a preference weighted measure. Health Technol Assess. 2012;16:1–165.

10. Al-Janabi H, Flynn T, Coast J. A self-report measure of capability well-being for adults: the ICECAP-A. Qual Life Res. 2012;21:167–76.

11. Flynn TN, Huynh E, Peters TJ, Al-Janabi H, Moody A, Clemens S, Coast J. Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. Health Econ. 2015. doi:10.1002/hec.3014 (online first).

12. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, Parkinson J, Secker J, Stewart-Brown SL. The Warwick–Edinburgh mental well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes. 2007;5(63). ISSN 1477-7525.

13. Office for National Statistics, Harmonised concepts and questions for social data sources: interim harmonised standard—personal well-being. 2014. http://www.ons.gov.uk/ons/guide-method/harmonisation/secondary-set-of-harmonised-concepts-and-questions/index.html. Accessed 23 Aug 2015.

14. Medical research Council. Improving cross-sector comparisons: beyond QALY. 2015. http://www.mrc.ac.uk/funding/how-we-fund-research/highlight-notices/improving-cross-sector-comparisons-beyond-qaly/. Accessed 23 Aug 2015.

15. Murray CJL, Vos T, Lozano R, Haghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2010;380:2197–223.

16. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, Begum N, Shah R, et al. Common values in assessing health outcomes form disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet. 2010;380:2129–43.

17. Gafni A, Birch S. Preferences for outcomes in economic evaluation: an economic approach to addressing economic problems. Soc Sci Med. 1995;40:767–76.

18. Brazier JE, Yang Y, Tsuchiya A, Rowen DL. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. Eur J Health Econ. 2010;11(2):215–25.

19. Gray A, Rivero-Arias O, Clarke PM. Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. Med Decis Making. 2006;26:18–29.

20. Hernández Alava M, Wailoo A, Ara R. Tails from the peak district: Adjusted Limited Dependent Variable Mixture Models of EQ-5D Health State Utility Values. Value Health. 2012;15(3):550–61. doi:10.1016/j.jval.2011.12.014.

21. Krabbe PFM, Stouthard MEA, Essink-Bot ML, Bonsel GJ. The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. J Clin Epidemiol. 1999;52(4):293–301.

22. Yang Y, Brazier J, Tsuchiya A. Effect of adding a sleep dimension to the EQ-5D descriptive system: A "bolt-on" experiment. Med Decis Making. 2013;34:42–53.

23. Dolan P, Kavetsos G, Tsuchiya A. Sick but satisfied: the impact of health and life satisfaction on choice between health scenarios. J Health Econ. 2013;32(4):708–14.

24. Brazier JE, Ratcliffe J, Tsuchiya A, Solomon J. Measuring and valuing health for economic evaluation. Oxford: Oxford University Press; 2007.

25. Rowen D, Brazier JE, Tsuchiya A, Hernandez M. Valuing states from multiple measures on the same VAS: A feasibility study. Health Econ. 2011;21(6):715–29.

26. Hernández Alava M, Brazier J, Rowen D, Tsuchiya A. Common scale valuations across different preference-based measures: estimation using rank data. Med Decis Making. 2013;33(6):839–52.

27. Dolan P, Kahneman D. Interpretations of utility and their implications for the valuation of health. Econ J. 2008;118(525):215–34.

28. McTaggart-Cowan H. Elicitation of informed general population health state utility values: a review of the literature. Value Health. 2011;14(8):1153–7.

29. Haybron DM. The pursuit of unhappiness: the elusive psychology of Well-being. Oxford: University Press; 2008.

30. Parfit D. Reasons and persons. Oxford: University Press; 1984.

31. Sen A. Commodities and capabilities. Amsterdam: North-Holland; 1985.

32. Nussbaum MC. Sex and Social justice. Oxford: University Press; 1999.

33. Anand P, Hunter G, Carter I, Dowding K, Guala F, van Hees M. The development of capability indicators. J Hum Dev Capab. 2009;10(1):125–52.

34. Burchardt T. Incomes, functionings and capabilities: the well-being of disabled people in Britain. PhD thesis, The London School of Economics and Political Science (LSE); 2005.

35. Dolan P, Lee H, Peasgood T. Losing sight of the wood for the trees: some issues in describing and valuing health, and another possible approach. Pharmacoeconomics. 2012;30(11):1035–49.

36. Mukuria C, Brazier J. Valuing the EQ-5D and the SF-6D health states using subjective well-being: a secondary analysis of patient data. Soc Sci Med. 2012;77:97–105.

37. Stewart-Brown S, Tennant A, Tennant R, Platt S, Parkinson J, Weich S. Internal construct validity of the Warwick–Edinburgh mental well-being scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey. Health Qual Life Outcomes. 2009;7:15.

38. Mukuria C, Rowen D, Peasgood T, Brazier J. An empirical comparison of well-being measures used in the UK. Policy Research Unit in economic evaluation of health and care interventions. Universities of Sheffield and York. EEPRU Research Report 027. Policy paper/document 01/04/2015; 2015.

39. Brazier J, Akehurst R, Brennan B, Dolan P, Claxton K, McCabe C, Sculpher M, Tsuchiya A. Should patients have a greater role in valuing health states? Appl Health Econ Policy. 2005;4(4):201–8.

40. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, Devlin N, Smith PC, Sculpher M. Methods for the estimation of the NICE cost effectiveness threshold. Centre for Health Economics, University of York, CHE research paper 81; 2013.

41. Claxton K, Walker S, Palmer S, Schulpher M. Appropriate perspectives for Health Care Decisions. CHE discussion paper 54, University of York, York, UK; 2010.

42. Department of Health. A new value-based approach to the pricing of branded medicines: a consultation. 2010. http://www.dhsspsni.gov.uk/medicineconsultation.pdf. Accessed 23 Aug 2015.

43. Donaldson C, Mason H, Shackely P. Contingent valuation in health care. In: Jones A, editor. The Elgar companion to health economics. Cheltenham: Edward Elgar; 2006.

44. Boadway RW, Bruce N. Welfare economics. Oxford: Basil Blackwell; 1984.

45. Dixon S, Shackley P. The use of willingness to pay to assess public preferences toward the fortification of foodstuffs with folic acid: results of a pilot study. Health Expect. 2003;6:140–8.

46. Olsen JA, Smith RD. Theory versus practice: a review of 'willingness-to-pay' in health and health care. Health Econ. 2001;10(1):39–52.

47. Donaldson C, Shackley P. Does "process utility" exist? A case study of willingness to pay for laparoscopic cholecystectomy. Soc Sci Med. 1997;44(5):699–707.

48. Robinson A, Gyrd-Hansen D, Bacon P, Baker R, Pennington M, Donaldson C. Estimating a WTP-based value of a QALY: the 'chained' approach. Soc Sci Med. 2013;92:92–104.

49. Gyrd-Hansen Kjaer, Gyrd-Hansen Dorte, Kjær Trine. Disentangling WTP per QALY data: different analytical approaches, different answers. Health Econ. 2012;21(3):222–37.

50. Kahneman D. Comments on the contingent valuation method. In: Cummings RG, Brookshire DS, Schulze WD, editors. Valuing environmental goods: a state of the arts assessment of the contingent valuation method. Totowa: Roweman and Allanheld; 1986. p. 18–194.

51. McFadden D, Leonard G. Issues in the contingent valuation of environmental goods: methodologies for data collection and analysis. In: Hausman JA, editor. Contingent valuation: a critical assessment. Amsterdam: North Holland; 1993.

52. Kahneman D, Ritov I, Schkade DA. Economic preferences or attitude expressions?: an analysis of dollar responses to public issues. J Risk Uncertain. 1999;19:203–35.

53. Skedgel CD, Wailoo AJ, Akehurst RL. Choosing vs. allocating: discrete choice experiments and constant-sum paired comparisons for the elicitation of societal preferences. Health Expect. 2013. doi:10.1111/hex.12098.

54. Ferrer-i-Carbonell A, van Praag BMS. The subjective costs of health losses due to chronic diseases. An alternative model for monetary appraisal. Health Econ. 2002;11(8):709–22.

55. Powdthavee N, van Den Berg B. 'Putting different price tags on the same health problem: re-evaluating the well-being valuation. J Health Econ. 2011;30(5):1032.