

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Baker, R; Chilton, S; Donaldson, C; Jones-Lee, M; Lancsar, E; Mason, H; Metcalf, H; Pennington, M; Wildman, J (2011) Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE. *Health economics, policy, and law*, 6 (4). pp. 435-447. ISSN 1744-1331 DOI: <https://doi.org/10.1017/S1744133111000181>

Downloaded from: <http://researchonline.lshtm.ac.uk/18589/>

DOI: [10.1017/S1744133111000181](https://doi.org/10.1017/S1744133111000181)

Usage Guidelines

Please refer to usage guidelines at <http://researchonline.lshtm.ac.uk/policies.html> or alternatively contact researchonline@lshtm.ac.uk.

Available under license: Copyright the publishers

Health Economics, Policy and Law

<http://journals.cambridge.org/HEP>

Additional services for *Health Economics, Policy and Law*:

Email alerts: [Click here](#)

Subscriptions: [Click here](#)

Commercial reprints: [Click here](#)

Terms of use : [Click here](#)



Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE

Rachel Baker, Sue Chilton, Cam Donaldson, Michael Jones-Lee, Emily Lancsar, Helen Mason, Hugh Metcalf, Mark Pennington and John Wildman

Health Economics, Policy and Law / Volume 6 / Issue 04 / September 2011, pp 435 - 447
DOI: 10.1017/S1744133111000181, Published online: 22 August 2011

Link to this article: http://journals.cambridge.org/abstract_S1744133111000181

How to cite this article:

Rachel Baker, Sue Chilton, Cam Donaldson, Michael Jones-Lee, Emily Lancsar, Helen Mason, Hugh Metcalf, Mark Pennington and John Wildman (2011). Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE. Health Economics, Policy and Law, 6, pp 435-447 doi:10.1017/S1744133111000181

Request Permissions : [Click here](#)

Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE

RACHEL BAKER

*Yunus Centre for Social Business and Health, Glasgow Caledonian University, Scotland, UK
Institute of Health and Society, Newcastle University, England, UK*

SUE CHILTON

*Institute of Health and Society, Newcastle University, England, UK
Newcastle University Business School, Newcastle University, England, UK*

CAM DONALDSON*

*Yunus Centre for Social Business and Health, Glasgow Caledonian University, Scotland, UK
Institute of Health and Society, Newcastle University, England, UK
Newcastle University Business School, Newcastle University, England, UK*

MICHAEL JONES-LEE

*Institute of Health and Society, Newcastle University, England, UK
Newcastle University Business School, Newcastle University, England, UK*

EMILY LANCSAR

*Institute of Health and Society, Newcastle University, England, UK
Newcastle University Business School, Newcastle University, England, UK
Centre for Health Economics, Monash University, Australia*

HELEN MASON

Institute of Health and Society, Newcastle University, England, UK

HUGH METCALF

*Institute of Health and Society, Newcastle University, England, UK
Newcastle University Business School, Newcastle University, England, UK*

MARK PENNINGTON

Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, University of London, England, UK

JOHN WILDMAN

*Institute of Health and Society, Newcastle University, England, UK
Newcastle University Business School, Newcastle University, England, UK*

Abstract: Recently, for many health economics researchers, empirical estimation of the monetary valuation of a quality-adjusted life year (QALY) has become an important endeavour. Different philosophical and practical approaches to this have emerged. On the one hand, there is a view that, with health-care budgets set centrally, decision-making bodies within the system can iterate, from

*Correspondence to: Cam Donaldson, Yunus Chair in Social Business and Health, NIHR Senior Investigator and Director, Yunus Centre for Social Business and Health, Institute of Applied Health Research, Glasgow Caledonian University, Level 3 Buchanan House, Cowcaddens Road, Glasgow G4 0BA, UK. Email: cam.donaldson@gcu.ac.uk

observation of a series of previous decisions, towards the value of a QALY, thus searching for such a value. Alternatively, and more consistent with the approach taken in other public sectors, individual members of the public are surveyed with the aim of directly eliciting a preference-based – also known as a willingness-to-pay-based (WTP-based) – value of a QALY. While the former is based on supply-side factors and the latter on demand, both in fact suffer from informational deficiencies. Sole reliance on either would necessitate an acceptance or accommodation of chronic inefficiencies in health-care resource allocation. On the basis of this observation, this paper makes the case that in order to approach optimal decision making in health-care provision, a framework incorporating and thus, to a degree, reconciling these two approaches is to be preferred.

Introduction

The question of the value to place on a quality-adjusted life year (QALY) has existed for some time in the health economics literature, notably since some authors sought to make the (controversial) link between cost-effectiveness analysis (CEA) and cost-benefit analysis (CBA; Phelps and Mushlin, 1991). More recently, the same question has come to the fore in policy circles through the creation of health technology assessment (HTA) agencies around the world. In making one-off recommendations about adoption or otherwise of an evaluated therapy, such agencies are, in effect, placing a monetary value on health gains. In the context of England and Wales, the National Institute for Health and Clinical Excellence (NICE) uses the QALY as its health metric. By operating with a threshold value of a QALY, above which an evaluated therapy is less likely to be recommended for adoption by the rest of the National Health Service (NHS), NICE is effectively setting the maximum price it will pay for, and thus the value it places on, a marginal QALY gained. Another way of looking at this is that the threshold represents NICE's estimate of the opportunity cost (or again the value of) services displaced at the margin by recommended technologies (National Institute for Health and Clinical Excellence, 2004).

Criticism of NICE as to the lack of an empirical basis for its current threshold has led to a recent and somewhat heated debate about how such a value should be arrived at. A natural starting point might be to adopt, at least in spirit, the standard economic approach to identify the value of any goods or services, which is to look at the equilibrium between demand and supply or, more specifically, the quantity (and associated price) at which marginal benefit equals marginal cost (MC). This is important, as it highlights the fact that for a 'true' value and hence optimum quantity of health care to be identified, decision makers such as NICE would in fact require information concerning both the supply and demand sides of the 'market'. Unfortunately, because no market exists for health care or QALYs, we cannot simply observe this value and instead research methods have been developed with the aim of eliciting the *value* or

cost of a QALY with a view to informing the NICE threshold.¹ Two main approaches have been advocated for obtaining a money value or cost per QALY, each essentially addressing the question from different sides of the ‘market’.

In the next section of this paper, we describe these two approaches in more detail and define those operating from the demand side as ‘surveyors’ and those from the supply side as ‘searchers’. While it may be tempting to regard these two perspectives as alternatives – or even rivals – in this paper we will argue that they are essentially complementary and that to rely exclusively on one or the other will, in the long run, almost certainly result in a sub-optimal allocation of society’s scarce resources. Thus, not only would we advocate a truce between searchers and surveyors, but would, indeed, suggest that both groups recognise that they are in fact fighting for the same cause! This should not be taken to imply that both should ultimately arrive at the same value for a QALY. We are agnostic as to how close or far apart they are or ‘should be’ and are not trying to impose an empirical relationship or validation of any theoretical link between NHS budget constraints and willingness-to-pay (WTP)-based values, as it is highly questionable whether such a link exists in theory.

After outlining what we mean by demand and supply in this context, we build on the theoretical framework put forward by Culyer *et al.* (2007) and present a formal exposition of the issues in an effort to outline such a reconciliation. Then, we address some recent criticisms levelled at surveyors’ methods and highlight the methodological challenges faced by both approaches. Such a response is necessary because not only is it important to address such criticisms but also, in doing so, issues emerge that direct us towards a forward-looking reconciliation of the perspectives.

Before proceeding, it is worth pointing out that there are also ongoing debates as to whether there is indeed a single value of a QALY ‘out there’ waiting to be discovered (Brouwer *et al.*, 2008) and whether we should weigh QALYs according to the characteristics of who is gaining them (Dolan *et al.*, 2008; Baker *et al.*, 2010). Similarly, we recognise that there are considerable challenges in estimating a WTP-based value of a QALY, as proposed by the surveyors (Pinto-Prades *et al.*, 2009; Baker *et al.*, 2010; Smith and Sach, 2010). Although these are clearly matters of considerable importance, this paper takes as its focus the recent discussion of issues of principle in arriving at some sort of ‘baseline’ value and the presentation of a formal, theoretical framework, which enables reconciliation of issues with respect to that.

1 It could be argued, of course, that markets exist for health-producing items such as over-the-counter (OTC) pharmaceuticals and that these promise QALY gains. However, challenges in arriving at a value of a QALY through data from such markets would face challenges at least equivalent to the contingent valuation methods we go on to discuss in this paper. Among others, the health gains bought are done so in very specific circumstances, not necessarily under conditions that reveal maximum willingness to pay, and many assumptions would have to be made about how the health gains being purchased translate into QALYs. Indeed, attempts at assessing ‘value’ in the OTC context often revert to contingent valuation methods (Lamiraud *et al.*, 2009). Nevertheless, it is important to acknowledge that other routes to the value(s) of a QALY exist.

Demand- and supply-side perspectives on valuing a QALY

The first approach, used by those we refer to as ‘surveyors’, is grounded solely on the demand side and focuses on the marginal benefits of a QALY. It assesses, through survey research, the WTP of members of the public for (typically small) health gains. These sums are then aggregated to arrive at an estimate of the overall value of a QALY. This approach is clearly in line with the general approach taken to CBA in the rest of the UK public sector, and, indeed, was used in a recent survey with respect to monetary valuation of a QALY funded by the Department of Health and undertaken in the United Kingdom (Baker *et al.*, 2010).

Those referred to in what follows as the ‘searchers’ have argued vigorously that bodies, such as NICE, should not elicit values from the public but should instead act as a ‘threshold-searcher’, seeking to identify the optimal threshold of incremental cost per QALY gained within the limits of expenditure set by the government (Culyer *et al.*, 2007; McCabe *et al.*, 2008). Hence, this approach can be thought of as being focused solely on the supply side of the ‘market’.² This search for a threshold would involve examination of past decisions taken by local NHS entities, such as primary care trusts (PCTs; Martin *et al.*, 2008; Appleby *et al.*, 2009). By identifying the price/cost of QALYs gained from interventions implemented and those lost by services displaced, the values of QALYs implied by such decisions can be inferred and compared with the NICE threshold, leading, presumably, to negotiations over what the ‘optimal’ value might be – although the process for determining optimality is unclear. In some senses, without NICE and PCTs negotiating over what the value of a QALY might be, it could be argued that the NICE threshold is just as problematic as a WTP-based one.

In building towards a conceptual framework, it is important to note here that we are not advocating a solely demand-side approach in which all health projects for which WTP is greater than the cost be implemented. In other parts of the public sector, imposed budget constraints mean that, in actual fact, WTP values are used to rank projects in order of net benefit. This requires knowledge of the benefit and cost curves faced by PCTs so as to maximise value from the budget, which implies that both demand- and supply-side approaches to the NICE threshold would be relevant.

Reconciling the issues between searching and surveying: a framework

Moving towards an alternative conceptual framework requires a shift in focus. Rather than viewing the two analytical perspectives as vying for supremacy, instead they should be seen as essentially complementary.

² By ‘supply side,’ we mean the whole of the NHS and setting to one side (for the purposes of this discussion) the internal or quasi-market in operation within the NHS in England.

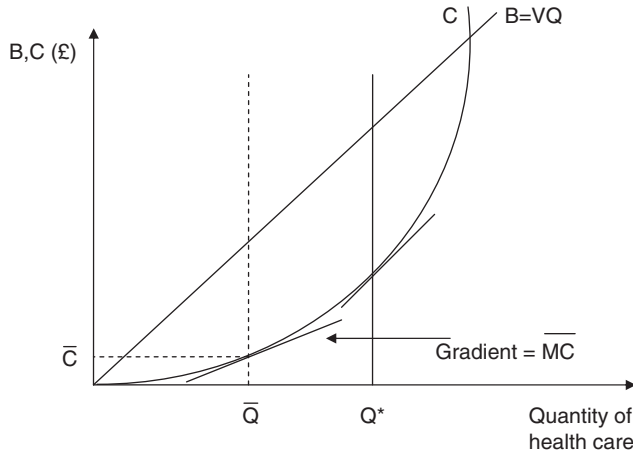


Figure 1. ‘Finding’ a value of a quality-adjusted life year.

By way of explanation, consider Figure 1 in which the quantity of health care, Q , is plotted on the horizontal axis and the corresponding total cost, C , of providing each overall level of health care – as well as the total benefit, B , associated with that level of provision – are plotted on the vertical axis.

By the standard type of welfare economics argument, provided that the value, V , per unit of health care (e.g. per QALY) has been defined on an aggregate WTP basis (possibly with distributional weights applied), then the total benefit associated with any level of provision, Q , will be given simply by VQ , so that the graph of total benefit (B) will be a straight line through the origin.³ In turn, on the assumption that the marginal cost of health-care provision is increasing, then the graph of the total cost function (C) will be increasing and convex. In the absence of any imposed public expenditure constraints, the optimal level of health-care provision will be at Q^* where marginal cost equals marginal benefit (or the gradient of $B =$ the gradient of C).

However, suppose that a constraint is imposed which requires that total expenditure on health-care provision during the period concerned should not exceed \bar{C} . The constrained optimum level of health-care provision for this period will then be given by \bar{Q} . In addition, suppose that health care is provided by a number of different health-care authorities (as is in fact typically the case) and that none of these authorities have full information concerning the total cost function. In spite of this, the constrained optimum will still be achievable provided that the central health-care authority permits the local authorities to

3 It might be argued that for large variations in Q it would be necessary to recognise that with diminishing marginal utility of health care, the graph of total benefit would be increasing and strictly concave. However, for more modest variations in Q of the type under consideration in this paper, the assumption of at least local linearity seems entirely reasonable. In any case, strict concavity of the total benefit function would not alter the essence of the argument that follows.

supply a service if and only if the marginal cost of doing so is less than or equal to the marginal cost, \overline{MC} , corresponding to \overline{Q} . This is essentially the searchers' approach. Alternatively, if the WTP-based value, V , per unit of health care is a multiple, k (>1), of \overline{MC} , then the local authorities could simply be required to provide a service only if its marginal benefit/cost ratio was greater than or equal to k . This is to all intents and purposes the essence of the surveyors' approach.

But, of course, in order to apply either of these two approaches, the central health-care authority must have adequate information concerning the nature of the cost function before issuing its guidelines to the local authorities. If it does not have this information at the beginning of the planning period, but instead faces a situation in which possible health-care projects will simply 'crop-up' at various stages of the period concerned, then the constrained optimum will simply be achievable only by chance. More specifically, without previous information concerning the properties of the total cost function it would be impossible to determine the critical 'cut-off' level, \overline{MC} , of marginal cost so that the searchers' approach would be a non-starter.⁴ Similarly, even if the magnitude of the WTP-based value, V , per unit of health-care service was well established, without previous knowledge of \overline{MC} it would be impossible to specify an appropriate cut-off benefit/cost ratio, k .

Hence without previous knowledge of \overline{MC} , the constrained optimum would, to all intents and purposes, be an unachievable ideal, both for searchers and surveyors. However, how then might the central health authority proceed? One possibility would be to rely on educated guesswork concerning the magnitude of \overline{MC} based on past experience, without any reference to the magnitude (or indeed the nature) of V . Alternatively, local health authorities could simply be informed about the level of V and instructed to consider only those proposed activities whose marginal cost of provision fell below this level and, so far as possible, to give priority to those activities with the lowest marginal cost.

Which of these two 'third-best' solutions should central health-care authorities adopt in practice? Our argument is essentially that, applied in isolation, each approach suffers from serious potential shortcomings but that if, by contrast, the two approaches are applied in parallel, as it were, then most of the more serious difficulties can be avoided.

As far as the potential shortcomings of each approach are concerned, suppose first that based on a 'search' of past decisions an educated guess is made concerning the appropriate level at which to set the cut-off marginal cost, \overline{MC} . Given that we are dealing with a situation in which the unconstrained optimum

⁴ At this point, it is worth noting a further, and potentially severe but generally overlooked, problem that is specific to the 'searchers' approach. As well as being based on imperfect information with respect to costs, a necessary assumption for implementation is that cost functions are stable over time and continuous. In effect, we assume that the past is a good proxy for the present and the future. This seems, at the very least, contestable. For example, inflation and innovation, among other things, will have had a significant impact on health-care provision costs. To ignore this feature is worrying.

is unachievable, it will necessarily be the case that \overline{MC} is less than V . However if V is not known (or at least information concerning the magnitude of V is not in the public domain), then it will be perfectly understandable if the press and media – and hence the general public – come to regard \overline{MC} as the ‘value’ that the central authority places on a unit of health care. The recent heated controversy over the implicit value placed on potential treatment for terminally ill cancer patients is a clear example. A further and arguably far more serious potential consequence of focusing exclusively on \overline{MC} and ignoring V is that central government decisions concerning the appropriate size of the overall health-care budget may be adversely influenced by the mistaken belief on the part of at least some key decision makers that \overline{MC} constitutes the appropriate measure of the marginal benefit of health care, with a resultant serious undervaluation of the latter in the overall allocation of public expenditure. Indeed, even if the size of the health-care budget is taken as given, there remains the possibility of a misallocation of the budget if the focus is exclusively on \overline{MC} . Thus, if – as many would argue ought to be the case – the overall benefits of health-care provision are defined sufficiently widely to include other effects such as time saving or the convenience of patients and their visitors (see e.g. Broome, 2004: 261), then there is clearly a danger that health improvement per se will tend to be underprovided relative to these other benefits if it is valued on the basis of \overline{MC} , without any reference to its ‘true’ value, V .

By contrast, suppose that no attempt is made to estimate \overline{MC} but that instead local health-care agencies are informed only about the level of V . While application of this value in the assessment of proposed health-care programmes will ensure that no treatments with a negative net social benefit will be undertaken, it is still possible that many programmes having a marginal cost in excess of \overline{MC} – and hence lying in the interval $[\underline{Q}, Q^*]$ in Figure 1 – will be (sub-optimally) undertaken.

However, if local health-care authorities are provided with a ‘best estimate’ of \overline{MC} and the level of V is clearly established and widely publicised, then the problems just outlined will be largely, if not entirely, avoided. These are precisely the grounds on which we would argue that the searchers’ and the surveyors’ approaches should be regarded as being essentially complementary rather than competing allocative tools.

Responding to the case against the surveyors

At this point, it would seem to be important to examine the case made against the use of WTP-based values in the health-care context, the demand-side concept advocated by the surveyors. While WTP-based values are now used extensively in public sector allocative and regulatory decision making (particularly in the context of transport, air pollution reduction and nuclear power generation), there are informational imperfections and potential weaknesses in the approach. Objections to the WTP approach would appear to revolve around three issues.

WTP and ability to pay

The first objection is that WTP is associated with ability to pay. It is then natural to think that defining the value of a QALY in terms of individual WTP would result in higher values for the better-off and lower values for those on lower incomes. This is not a criticism that has been made by recent searchers, but is rather a more impressionistic and common criticism, which nevertheless finds its way into the more formal literature from time to time. A good example is the following quote from the authors of the Washington Panel Guidelines on Cost Effectiveness:

Cost-benefit analysis's primary valuation method is willingness to pay (WTP), an approach whose difficulty lies in its intrinsic favouring of the programs and diseases of the affluent over those of the poor.

(Gold *et al.*, 1996)

The searchers' solution to this is to have decisions made on the basis of QALYs gained for resources invested or equivalently, cost per QALY-gained, the implication being that such an approach is value free with respect to unpleasant distributional considerations. This strategy, however, seems to ignore the fact that it has also been shown that QALYs are subject to the same distributional challenges as WTP (Donaldson *et al.*, 2002), the solution for which is the same: to average stated utility values across populations.

Nevertheless, it is incumbent on surveyors to address this issue specifically in respect of a WTP-based value. Noting, as above, distributional problems associated with a QALY-based value may be necessary, but it is not sufficient. To be clear, the aim of asking a cross-section of the population about their WTP is to obtain a value that takes proper account of the preferences and wishes of members of the public, as well as society's overall budget constraint. Advocates of the WTP approach, the surveyors, would tend to argue for one value, based on some measure of central tendency, to be applied to each member of society regardless of income. Indeed, public sector agencies that use such values (for example, the Department for Transport) invariably do apply the *same* value to all income groups once that value has been calculated by averaging across the population. Recent advances in the safety economic literature have provided an economic theoretic foundation for this previously *ad hoc*, but arguably fair, approach to policymaking (Baker *et al.*, 2008). Such a measure of central tendency is crucial to the more formal framework set out above.

Detachment from the budgetary process

The second argument, and the main one made by the searchers, is that, as the NHS budget is set through negotiation between the Treasury and the Department of Health, individual WTP values have no relevance for health service resource allocation as they are somehow detached from the budget-setting process. This has led to the stated view that

information about how much an individual or society values improvements in health (i.e. their willingness to pay for a QALY) is not at all relevant to the NICE remit.

(Culyer *et al.*, 2007, p. 57)

This is a strong argument, but it can be countered in three ways. The first and most important point is that in a democratic society there is a powerful case to be made for ensuring that the government's budget-setting process should, so far as possible, be fully informed about the preferences and attitudes of members of the public. Thus, while it is reasonable to expect that negotiations between the Treasury and the Department of Health will take account of a number of different factors, information concerning the public's WTP for health care should arguably constitute an important consideration, the role that information could play having been outlined more formally above.

Second, NICE can be subjected to the same criticism of being detached from the NHS budget as that levelled at the surveyors. NICE does not have a budget, and PCTs constantly complain about decisions being imposed upon them from NICE. Such complaints, if valid, would mean that the current threshold is leading to a misallocation of NHS resources which, ultimately, means we are getting less health for our tax £s spent.

A third consideration is that, while surveyors' values may be detached from the budgetary process, searcher values are in fact *too* attached. There is a self-fulfilling prophecy (more formally, path-dependency) to the process whereby NICE searches for a threshold through a series of decisions it has already made. One would predict that the value to emerge from this process is likely to be around the threshold range of £20k–30k per QALY, precisely because this is the value at which NICE already operates! The literature abounds with examples of decision-makers' judgements leading to great variation in values of lives saved (Mooney, 1984). It may well be that £20k–30k is the correct range because it was made on the basis of intelligent guesses by experts at NICE's outset. However, we can never know whether this is the case or whether there is a path dependency towards the range because that is what is now 'out there'. Indeed, the influence of the stated threshold range on decisions has already been demonstrated in the literature (Devlin and Parkin, 2004).

A threat to the Constitution?

The WTP approach has been depicted as a threat to the constitutional jurisdiction of Parliament! The argument put forward by threshold-searchers is that survey-based WTP approaches somehow substitute the 'direct democracy' of public opinion for the parliamentary process through which the Treasury and Department of Health agree a budget for the NHS, within which PCTs and NICE have to work (Culyer *et al.*, 2007). It is claimed that 'experimental methods' (i.e. WTP studies) cannot capture the opportunity costs more effectively than Parliament as the budget-setting process for the NHS takes account

of assessments of the marginal value of extensions of a wide range of public programmes and of the value of purchasing power left in the pockets of consumers (McCabe *et al.*, 2008).

The simple answer to this is to ask what can be more legitimate than seeking to determine what those who stand to gain or lose from different allocations of NHS resources – that is, the public – think should count, not least as these resources are ultimately paid for by the public, that is taxpayers, themselves. (NICE does this anyway, as the quality-of-life ‘tariff’ that it uses to calculate the cost per QALY of interventions is based on a survey of the general public.) We would also contend that, beyond a single value of a QALY, information is needed regarding the relative value of QALYs made up largely of improvements in quality of life vis-à-vis those emanating from survival gains, and that the only legitimate way of doing this (and of avoiding politicians and decision-makers being criticised for the decisions made) is to establish more thoroughly the preferences of the public as regards different types of health gain – an issue that has grown in importance with the emergence of value-based pricing in England as the bulk of this paper was written. As we have already noted, we do not regard information concerning the public’s WTP for health care as being the only relevant input in the NHS allocative decision-making process, but we do believe that it constitutes an important consideration. Furthermore, if, in its budget-setting, Parliament has already taken account of the purchasing power left in the pockets of consumers, the gap between a value of a QALY inferred from WTP surveys of the public and from NHS decisions might not be that great.

Outstanding methodological challenges in developing a new framework for QALY valuation

Having digressed somewhat to defend the use of WTP-based values in health-care decision making, it is important to return to our main objective of reconciliation. Both surveyors and searchers face significant challenges in designing and executing research that delivers reliable demand and supply side information. For surveyors, these challenges concentrate around issues that have been discussed for some time in the economics literature, such as overcoming framing effects and insensitivity of responses to small changes in scale, and which have also come to light in recent value-of-a-QALY research (Pinto-Prades *et al.*, 2009). Methodological issues that are perhaps more specific to the pursuit of WTP per QALY include questions regarding the aggregation of individual ratios (Baker *et al.*, 2010), appropriate measures of central tendency and the effect on elicited values of ‘non-traders’ in either health or wealth domains.

Threshold-searchers face different challenges. Examining investment and disinvestment decisions made by PCTs in order to establish the cost per QALY at the margin is one way of searching for a threshold value. If, for example,

disinvestments made by PCTs were associated with a lower cost per QALY than recommendations made by NICE, it would imply that the threshold is too high. However, the idea that PCTs can somehow iterate towards such a value would seem to be rather farfetched at the moment. To date, in the NHS, no systematic framework for commissioning, which recognises scarcity and can explicitly address tradeoffs, has been implemented. We know this as health economists who have done more than most in terms of working with the NHS on trying to improve its decision-making processes (Bate *et al.*, 2007). It would seem that the development of such a framework is essential for matching national priorities with local needs and to provide local health organisations with a defensible mechanism for (occasionally) justifying a focus on the local as well as the national agenda. When in place, perhaps we could iterate towards the value of a QALY at the margin. However, only a few years ago, the Health Committee of the UK House of Commons stated that practical systems and structures should be put in place to improve capacity to implement guidance, as implicit prioritisation is insufficient, and the Government must work towards “a comprehensive framework for health-care prioritisation, underpinned by an explicit set of ethical and rational values to allow the relative costs and benefits of different areas of NHS spending to be comparatively assessed in an informed way”. The more recent assessment of PCTs, through the World Class Commissioning initiative, would seem to show little movement towards this ideal.

Appleby *et al.* (2009), in their recent and admirable quest, state that “a definitive finding about the consistency or otherwise of NICE and NHS cost-effectiveness thresholds would require very many decisions to be observed, combined with a detailed understanding of the local decision-making processes”. The most rigorous valuations from the searchers to date are those of Martin *et al.* (2008) who attempted to estimate what it costs the NHS in England to produce a QALY in the programme areas of cancer and circulatory disease; arriving at values of £11,960 and £19,070, respectively. However, the authors recognise that these can only be regarded as very rough estimates and the availability of high-quality data on which to base estimations of QALYs (as opposed to mortality estimates) is limited.

Even if robust cost per QALY figures could be estimated for all 23 programme areas within the National Programme Budgeting Dataset for England, it is likely that costs will vary widely. It is also well known that costs for some conditions, such as heroic treatment for neuroblastoma, are in excess of £120,000 per QALY gained. Identification of such a range is useful, but it is not clear how much further, compared with other approaches, it gets us towards validating or querying the NICE threshold. Strictly, they tell us the costs (and not value) of producing a QALY in different treatment circumstances. If some of these costs are as high as for neuroblastoma, information on the public’s WTP for a QALY, both in absolute terms and (as yet, little researched) on how WTP across the range of more-heroic to more-routine scenarios varies, would be useful.

Given the uncertainties involved in empirical approaches based on all the methods outlined above, it would seem sensible to articulate all of them within one framework that can show how a value can be ‘triangulated’ from the NICE threshold, NHS costs of producing a QALY and the public’s WTP for a QALY.

What is needed now?

In order for a framework based on the two approaches to be developed, it requires acceptance on each ‘side’ that information deficiencies will continue to persist, both in respect of supply and demand factors; to rely solely on one method almost certainly delivers inefficient health-care resource allocation. If this premise is accepted, then there is a chance to progress towards a framework that embraces both, maximising the information available to decision makers. Many other considerations will of course be overlaid on this basic position but it seems a reasonable place to start given the highly challenging question being addressed.

Practically, as results become available from further refined surveys of the public’s WTP and improved observations of NICE and PCT decisions, we can then begin to speak in more informed ways about the size of the $Q^* - \bar{Q}$ gap. As we have stated, we are agnostic about its magnitude, but we noted above that the number(s) arising from the surveyors’ research may not be as far from the search-based ones as might be speculated.

Despite the work that has been started, what is now required is more in-depth survey research on the public’s views on the value they might attach to the different ways QALYs can be generated and on the relative importance that might be given to the characteristics of potential beneficiaries of NHS services. Research to encourage improved and more systematic decision making by PCTs would also help as would that using routinely collected data on expenditures and outcomes, and, eventually, would permit reasoned comparison with the views of the public. Then, as well as observing the size of the $Q^* - \bar{Q}$ gap, we can also have a more reasoned debate on what, if anything, to do about it.

References

- Appleby, J., N. Devlin, D. Parkin, K. Chalkidou and M. Buxton (2009), Searching for cost effectiveness thresholds in the NHS, *Health Policy*, 91: 239–245.
- Baker, R., S. Chilton, M. Jones-Lee and H. Metcalf (2008), Valuing lives equally: defensible premise or unwarranted compromise?, *Journal of Risk and Uncertainty*, 36: 125–138.
- Baker, R., I. Bateman, C. Donaldson, M. Jones-Lee, E. Lancsar, G. Loomes, H. Mason, M. Odejar, J.-L. Pinto Prades, A. Robinson, M. Ryan, P. Shackley, R. Smith, R. Sugden and J. Wildman (2010), Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY project, *Health Technology Assessment*, 14(27).

- Bate, A., C. Donaldson and M. Murtagh (2007), Managing to manage scarce resources in the English NHS. What can economics teach? What can economics learn?, *Health Policy*, **84**: 249–261.
- Broome, J. (2004), *Weighing Lives*, Oxford: Oxford University Press.
- Brouwer, W., J. van Exel, R. Baker and C. Donaldson (2008), The new myth: the social value of the QALY, *Pharmacoeconomics*, **26**(1): 1–4.
- Culyer, A., C. McCabe, A. Briggs, K. Claxton, M. Buxton, R. Akehurst, M. Sculpher and J. Brazier (2007), Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical excellence, *Journal of Health Services Research and Policy*, **12**: 56–58.
- Devlin, N. and D. Parkin (2004), Does NICE have a cost-effectiveness threshold and what other factors influence its decisions?, *Health Economics*, **13**: 437–452.
- Dolan, P., R. Edlin, A. Tsuchiya, C. Armitage, A. Hukin, J. Brazier, R. Ibbotson, S. Bryan, D. Eiser and J. S. Olsen (2008), The relative societal value of health gains to different beneficiaries. Report for the NIHR Methodology Programme, January.
- Donaldson, C., S. Birch and A. Gafni (2002), The pervasiveness of the ‘distribution problem’ in economic evaluation in health care, *Health Economics*, **11**: 55–70.
- Gold, M. R., J. E. Siegel, L. B. Russell and M. C. Weinstein (1996), *Cost-effectiveness in Health and Medicine*, New York: Oxford University Press.
- Lamiraud, K., K. Von Bremen and C. Donaldson (2009), The impact of information on patient preferences in different delivery patterns: prescription versus OTC, *Health Policy*, **93**: 102–110.
- Martin, S., N. Rice and P. C. Smith (2008), Does health care spending improve health outcomes? Evidence from English programme budgeting data, *Journal of Health Economics*, **27**: 826–842.
- McCabe, C., K. Claxton and A. J. Culyer (2008), The NICE cost-effectiveness threshold: what it is and what it means, *Pharmacoeconomics*, **26**: 733–744.
- Mooney, G. (1984). *Valuing Human Life in Health Service Policy*. Nuffield/York Portfolios No. 3, Nuffield Trust, London.
- National Institute for Health and Clinical Excellence (2004), *Guide to the Methods of Technology Appraisal*, London: National Institute for Health and Clinical Excellence.
- Phelps, C. E. and A. I. Mushlin (1991), On the (near) equivalence of cost-effectiveness and cost–benefit analyses, *International Journal of Technology Assessment in Health Care*, **7**: 12–21.
- Pinto-Prades, J. L., G. Loomes and R. Brey (2009), Trying to estimate a monetary value for the QALY, *Journal of Health Economics*, **28**: 553–562.
- Smith, R. and T. Sach (2010), Contingent valuation: what needs to be done?, *Health Economics, Policy and Law*, **5**: 91–111.