



QALYs and the capability approach

Richard Cookson*

School of Medicine, Health Policy and Practice, University of East Anglia, UK

Summary

This explores the applicability of Sen's capability approach to the economic evaluation of health care programmes. An individual's 'capability set' describes his freedom to choose valuable activities and states of being ('functionings'). Direct estimation and valuation of capability sets is not feasible at present. Standard preference-based methods such as willingness to pay are feasible, but problematic due to the adaptive and constructed nature of individual preferences over time and under uncertainty. An alternative is to re-interpret the QALY as a cardinal and interpersonally comparable index of the value of the individual's capability set. This approach has limitations, since the link between QALYs and capabilities is not straightforward. Nevertheless, the QALY approach is recognisable as an application of the capability approach since it pays close attention to functionings, through the use of survey-based multi-attribute health state valuation instruments, and permits conceptions of value other than the traditional utilitarian ones of choice, desire-fulfilment and happiness. Furthermore, suitably re-interpreted, it can account for (i) non-separability between health and non-health components of value; and suitably modified it can also account for (ii) process attributes of care, which may have a direct effect on non-health functionings such as comfort and dignity, and (iii) sub-group diversity in the value of the same health functionings. Copyright © 2005 John Wiley & Sons, Ltd.

Keywords capability approach; economic evaluation; welfarism; QALY; process attributes

Introduction

Sen's capability approach assesses individual well-being in terms of 'capability sets' that describe what individuals are free to do or to be [1–4]. This approach has philosophical pedigree and is recognised as the leading alternative to standard welfare economic theory [5]. It has engaged a diverse range of academic disciplines and policy communities, and has been extensively applied in the area of development [6]. It is therefore natural for health economists to ask: how can this approach be applied to economic evaluation of health care programmes?

I shall assume that, to answer this question, we need to find a way of making (partial) cardinal and

interpersonal comparisons of the value of individual capability sets. This is because we need to weigh gains for some people (e.g. patients who benefit from a cost-increasing new health care programme) against losses for others (e.g. other patients or taxpayers who bear the opportunity costs of the new programme through reduced quality of care or income). Here I part company from some economists, who believe that cardinal interpersonal comparisons are unnecessary or unwarranted [5,7].

I shall further assume that such comparisons need to be *quantified*. Here, I part company from some Sen scholars, who warn against 'systematically narrowing the approach down to a technical tool or an algorithm that 'measures' non-quantifiable dimensions' [6].

*Correspondence to: Centre for Health Economics, University of York, York YO10 5DD, UK. E-mail: rc503@york.ac.uk, r.cookson@uea.ac.uk

Three ways of applying the capability approach to health economic evaluation are examined: (1) direct estimation and valuation of capability sets, (2) standard 'preference-based' methods such as willingness to pay, and (3) a re-interpretation of the QALY approach. The first approach is not feasible at present. The second is feasible, but argued to be inadequate for the task due to the adaptive and constructed nature of individual preferences over time and under uncertainty. The third approach suffers from a number of limitations, since the link between QALYs and capabilities is not straightforward. Nevertheless, the QALY may be useful in health care contexts as an imperfect but workable index of the value of the individual's capability set.

My proposal is to re-interpret the QALY as a cardinal and interpersonally comparable index of the value of an individual's capability set in a given time period under certainty. This differs from the more usual interpretations of the QALY found in the literature. For example, some authors interpret the QALY as an index of von-Neumann–Morgenstern expected utility – a normative representation of the individual's preferences between uncertain outcomes, incorporating attitude towards risk [8]. By contrast, other authors interpret the QALY as a cardinal and interpersonally comparable index of health [9,10]. We might call these three interpretations the 'capability QALY', the 'utility QALY' and the 'health QALY', respectively.

The paper is structured as follows. The next section summarises Sen's capability approach and relates it to the 'welfarist' versus 'extra-welfarist' debate in health economics. The section following this explores some of the difficulties with direct estimation and valuation of capability sets. The succeeding section examines the possibility of valuing capability sets using standard preference-based methods such as willingness to pay. The next section sets out my proposed QALY approach in relation to one of its key features: the imposition of a simple additive structure of value across time and states of nature. The section that follows explores a second key feature: the value of full health is assumed to be the same for all individuals. The penultimate section analyses the differences between the 'health QALY' and the 'capability QALY', using some simple notation, and identifies three main advantages of the latter: (i) it allows non-separability between health and non-health components of the value of capability sets, (ii) it can account for the value of process

attributes of care and (iii) it can account more easily for subgroup diversity in the value of health states. The last section concludes.

Background: the capability approach

Over the last quarter century, Sen has developed a respected approach to assessing individual well-being that he calls the capability approach [1–4,11]. The approach is now central in the development field, and has formed the conceptual framework for all United Nations Human Development Reports since 1990 [12]. It is also making inroads elsewhere and has generated a large, rapidly expanding and cross-disciplinary literature. For a comprehensive and up-to-date bibliography of this literature, including recent attempts to apply the capability approach across a range of policy sectors, see www.fas.harvard.edu/~freedoms/bibliography.html.

It is hard to pin down a clear and definitive statement of the capability approach. The approach has developed gradually through numerous publications written for different disciplinary audiences at different times and in different policy sectors [6]. Sen seems to say different things at different times – and sometimes conflicting interpretations are possible. In what follows, however, I attempt to identify certain key features of the approach that contrast with standard approaches to assessing individual advantage within economics.

The capability approach was originally developed for the purpose of evaluating inequality, as an answer to the question 'equality of what?'. As Sen says, however, 'as well as the question "equality of what?" there is the parallel question "efficiency of what?"' [3]. In principle, Sen's approach can thus be applied to efficiency questions – including questions about the economic evaluation of health care programmes.

Sen sees his capability approach as an alternative to the standard 'welfarist' approaches to individual advantage that dominated welfare economics and utilitarian philosophy during the 20th century [5,11]. 'Welfarist' approaches assess states of affairs in terms of individual welfare or utility. By contrast, the capability approach assesses states of affairs in terms of the individual's freedom to pursue valuable acts or reach valuable states of being. It can therefore be thought of as an

'extra-welfarist' approach, in the broad sense that individual utility is not necessarily the only focus of attention in assessing states of affairs [9,13,14]. However, the capability approach is not 'extra-welfarist' in any narrow sense of assessing states of affairs exclusively in terms of the health functionings of individuals while ignoring non-health functionings [15].

To avoid confusion, it is necessary further to clarify the terms 'utility' and 'welfarism', which mean different things to different people. As Sen says: 'the term 'utility' is often used to mean quite different things, and there is a new – but by now widely used – tradition by which anything of value is called by that versatile name' [2, p. 12].

We can distinguish at least three different interpretations of 'welfarism' that depend on three different interpretations of 'utility'. First, there is the narrow interpretation that sees 'welfarism' as largely synonymous with standard welfare economic theory. Standard welfare theory adopts an ordinal choice-based interpretation of utility, often restricting the domain of choice to the commodity space [16,17]. This narrow interpretation has been a target of criticism by numerous authors of an 'extra-welfarist' persuasion – including Culyer, who coined the phrase [9]. Second, there is Sen's own interpretation of 'welfarism'. This allows, in addition to the choice-based interpretation of utility used in standard welfare economics, various other standard interpretations of utility within the utilitarian philosophical tradition [2, pp. 12–13]. These other interpretations of utility include the 'happiness' view, the 'desire-fulfilment' view, and variations thereof [18]. Sen sees his capability approach as a departure from 'welfarism' defined in this way. Finally, there is the broad interpretation that allows utility to represent anything of value. If 'utility' and 'welfarism' are defined as broadly as that, then Sen's capability approach can be seen as 'welfarist' – as some health economists have claimed [15].

Sen also rejects definitions of individual advantage in terms of income or resources or access to goods and services. One reason is that people differ radically in their ability to convert resources into valuable activities and states of being, due to diversity in people's internal characteristics (e.g. health, strength, stamina, charisma) and external circumstances (e.g. location, social position, employment, family circumstances).

The building blocks of the capability approach are a set of valuable dimensions of wellbeing –

what Sen calls 'functionings'. These might range, for example, from elementary functionings such as mobility through to more sophisticated functionings such as the ability to go scuba diving or to feel a sense of accomplishment. What Sen calls 'elementary evaluation' involves constructing an index of achieved functioning. This is done by giving each functioning a score and then combining them using a set of dimensional weights. That is how, for example, the UN Human Development Index is constructed.

Sen further distinguishes the individual's actual level of functioning ('achieved wellbeing') from his ability to achieve different levels of functioning ('wellbeing freedom'). The latter is described by the individual's 'capability set' – the set of functionings the individual is capable of achieving. Sen refers to wellbeing freedom as a concept of 'effective freedom' (freedom to achieve valuable outcomes), to distinguish it from the standard philosophical concepts of 'negative freedom' (freedom from coercion by others) and 'positive freedom' (self-mastery) [19,20]. Effective freedom to choose is intrinsically valuable, since an autonomous life is better than a controlled one. For instance, it may be in a person's interests to have the ability to vote in free and fair elections even if they do not actually vote. Effective freedom may also be instrumentally valuable, since the exercising of choice may foster the kinds of learning and self-improvement that lead, ultimately, to improved functioning.

To clarify the difference between 'achieved wellbeing' and 'wellbeing freedom', consider Figure 1. This illustrates a simple capability set

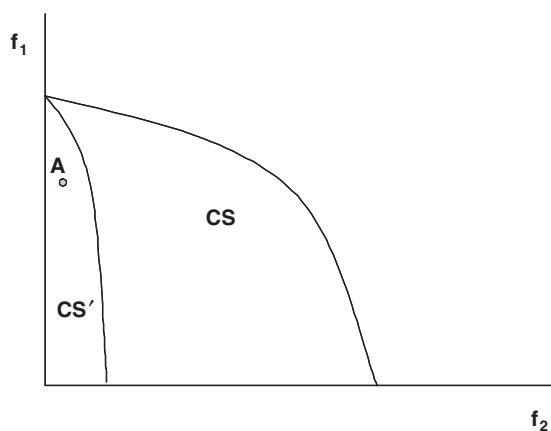


Figure 1. Capability sets – a simple illustration

with two functionings, f_1 and f_2 , for a single individual. Imagine that f_2 is an index of the individual's level of 'political participation' and f_1 a composite index of all other functionings. CS describes the individual's current capability set, and point A represents the individual's chosen level of functioning. This individual has a low level of political participation – he never bothers to vote or to engage in political protests. CS' describes an otherwise identical situation except that various political rights have been abolished – e.g. voting rights, freedom of assembly and so on. The individual's level of achieved functioning, A , does not change. So his 'achieved wellbeing' – i.e. the value of A – remains the same. However, the individual's *freedom* to participate in politics has been limited. So his 'wellbeing freedom' – i.e. the value of his capability set as a whole – may have fallen.

Note that the capability approach differs from standard welfare economic evaluation in one other important respect, since it relaxes the assumption of rational self-interest [21]. The individual does not necessarily choose the 'best' point in his capability set (i.e. the one that best serves his own interests). So A does not have to lie on the functioning possibility frontier.

Against critics who argue that his approach requires too much value judgement on the part of the analyst [5], Sen replies that 'the need for selection and discrimination is neither an embarrassment, nor a unique difficulty, for the capability approach' [3]. In other words, all methods of economic evaluation allow the analyst leeway to select the assumptions and evidence that go into the analysis. This selection involves implicit value judgements. So, according to Sen, the capability approach does not ultimately involve 'more' value judgement than rival approaches; rather, it forces the analyst to make those value judgements explicit. This can be seen as an advantage, since one of the main justifications for doing economic evaluation is that it facilitates open government [22].

Sen emphasises two further methodological points about his approach. First, evaluation in the space of capabilities is not the only relevant information for decision-making. Wider considerations such as procedural rights and duties may also matter – and not just because they influence people's capabilities. Second, capability evaluation should be seen as a partial exercise that does not permit complete rankings of all social states. In a

pluralistic society, ambiguity and disagreement will arise at all stages of capability evaluation – in the selection of functionings, in dimensional scoring and weighting procedures, in estimating and valuing individual capability sets, and in the formula for aggregating individual values to yield social rankings. Sen argues that, 'if an underlying idea has an essential ambiguity, a *precise* formulation of that idea must try to *capture* that ambiguity rather than lose it [11, p. 49]'. (Italics in original.)

Direct methods of applying the capability approach

Unfortunately, it is not possible at present to directly apply the capability approach to economic evaluation. One problem is that it is hard to secure agreement about the appropriate list of functionings. Allowing an indefinitely long list of functionings cannot solve this problem. The longer the list of functionings, the harder it becomes to establish coherent trade offs between them using responses to value elicitation questions. This is for the 'bounded rationality' reason that respondents tend to simplify questions as they can only deal with a few aspects of the situation at a time [23].

The move from functioning to capability is even more problematic. In theory, it would be possible to estimate capability sets, like any other opportunity sets, through use of a 'reference group' approach. This approach has been used for example to examine how far non-employment can be considered a voluntary choice [24]. The basic idea is to investigate the actual choices made by reference groups of individuals who share similar observable characteristics that may act as constraints on their opportunity sets (e.g. age, gender, education level, location, caring responsibilities, and so on). If there is a high probability that an individual in a particular reference group chooses a particular functioning (e.g. employment), then it is assumed that this functioning lies within his capability set (whether or not he actually chooses it). Differing views about what factors are regarded as beyond the individual's control can be accommodated by using different characteristics to define the reference group.

One difficulty with this approach is potential confounding between opportunities and ‘group preferences’: individuals in the reference group may not choose something because most of them do not want it, not because it is not available to them. However, it might be possible to shed light on this difficulty using a survey of the general population to identify what is regarded as ‘normal’ functioning. If a functioning is ‘normal’, but the individual has not achieved it, then this may reasonably be attributed to a constraint rather than a choice. Another difficulty is that this approach is data-hungry: it requires large samples of detailed individual level data linking achieved functionings to individual characteristics. Estimating complex capability sets involving multiple functionings each with multiple levels may therefore not currently be feasible.

Even if complex capability sets could be estimated, however, there is no agreement on how to compare them – except, of course, for the special case where one capability set is fully contained within and thus dominated by another. Difficulties arise in trading off different valuable attributes of the set, such as its diversity, expected value, maximum value and so on. Sugden has investigated this general problem in relation to comparing opportunity sets (of which capability sets are a special case) and has concluded that all methods are flawed [20]. He argues that the most promising approach involves the use of ‘potential preferences’. The idea is that an opportunity set is valuable to an individual insofar as it caters to the range of potential preferences the individual might have had. The method Sugden proposes for measuring potential preferences – what he calls the ‘sociological method’ – involves investigating the choices of a reference group of individuals. The probability that an individual might have had a particular preference ordering is the frequency with which that ordering occurs within his reference group.

However, just because a particular functioning is popular among your reference group does not necessarily guarantee it is valuable to you, and *vice versa*. For example, the ability to play a musical instrument may be valuable to a child, even if almost all of his peers prefer to play computer games. Furthermore, use of ‘reference group’ approaches both to estimate capability sets and to compare them might result in a problem of confounding. Do the observed choices of your reference group indicate the shape of your oppor-

tunity set (as required for estimating sets), or your potential preferences between opportunity sets (as required for comparing sets), or a bit of both?

Sugden’s analysis of opportunity metrics concludes that: ‘Perhaps the most we can expect to find are imperfect but workable indices of opportunity’ [20]. We now turn to one proposal for an imperfect but workable index of capability – willingness to pay.

Standard ‘preference-based’ methods such as willingness to pay

It might be possible to adapt standard ‘preference-based’ economic valuation methods such as willingness to pay surveys to the task of valuing capability sets. One method, for example, might be to estimate how much each individual is willing to pay for specified improvements in their capability set brought about by a particular health intervention. Some health economists appear optimistic about the idea of using preferences to value capability sets, even going so far as to suggest that ‘under Sen’s approach... preferences remain paramount’ [15].

It is important, however, to be clear about what is meant by ‘preferences’. Standard ‘preference-based’ economic valuation methods (both ‘revealed preference’ and ‘expressed preference’ methods) gather data about people’s choices and/or desires. However, Sen explicitly rejects the use of either choices or desires to value capabilities:

‘The choice approach to well-being is...really a non-starter’ [2, p. 14].

‘If, on the other hand, desire-fulfilment is taken as the criterion, then a very particular method of evaluating capabilities and functionings would have been chosen. The adequacy of this particular perspective for the evaluation of capabilities and functionings is deeply disputable, since any mechanical use of a metric of desires rather than facing the problem of reasoned assessment does injustice to the exercise of normative evaluation’ [11, pp. 54–55].

We can distinguish people’s choices and desires from their value judgements. Choices and desires are ‘positive’ descriptions of the world – either of individual behaviour or a mental state. It is largely up to the individual what conscious or unconscious motivations lie behind his choices and

desires; analysts and fellow citizens generally have little or no business investigating those motivations or engaging the individual in debate about their rights and wrongs. By contrast, value judgements are normative and by nature susceptible to reasoned assessment.

Sen's approach is based on value judgements, rather than choices or desires. Furthermore, it is based on value judgements of a particular kind: the reasons given for those value judgements must ultimately relate to the individual's capability set. Insofar as individual choices or desires do not relate to the individual's capability set, they are not relevant to capability evaluation. Like choices and desires, however, value judgements can be elicited using opinion surveys – and in this rather loose sense can be referred to as 'preferences'. So responses to 'willingness to pay' questions could in principle be used to value capability sets – but only if they are interpreted as value judgements, rather than desires or choices.

This represents a radical departure from standard welfare economic theory. Standard theory insists that 'willingness to pay' amounts are a money metric representation of *ordinal non-comparable* utilities that represent individual behaviour [16,17]. Standard welfare theory does not permit direct interpersonal comparisons when there are both winners and losers from a policy. Instead, it uses a compensation test. For example, the Hicks–Kaldor compensation criterion allows one to identify a 'potential Pareto improvement' – although only under certain conditions [25]. It is possible to re-interpret willingness to pay amounts as cardinal and interpersonally comparable measures of value – but such a re-interpretation has no basis in standard welfare economic theory.

It is also important to distinguish between using preferences as the *source* of value versus as *evidence* for value [2]. Sen's approach explicitly rejects the former, but not the latter. If an individual prefers capability set *A* to *B*, this may be strong *prima facie* evidence that *A* is more valuable to him than *B*. According to Sen's approach, however, it is nevertheless logically possible that *B* is more valuable to him than *A*. It is therefore possible to apply Sen's approach using 'preference-based' (or, more accurately, 'survey-based') methods – but only in the weak sense that responses to survey questions may be used as evidence for value, not in the strong sense that value is defined in terms of individual choices or desires.

One reason why Sen cautions against using preferences to define value is that they may adapt to circumstances. Individual choices, desires and judgements all depend on expectations. Disadvantaged individuals may lower their expectations (the problem of 'entrenched deprivation') while advantaged individuals may raise their expectations (the problem of 'expensive tastes'). One way that deprived (or unhealthy) individuals cope with adversity is by lowering their expectations. By doing so, they may achieve greater desire-fulfilment than well-to-do individuals with high expectations, thus, removing any argument for redistributing resources towards the former. Such problems may be particularly pressing when making international comparisons between developed and developing countries. As Sen points out, for example, self-reported morbidity statistics can yield misleading comparisons between developed and developing countries since they are influenced by education levels, availability of health care facilities, and public information on illness and remedy [26].

Another, and perhaps more pressing problem in the context of health care evaluation in developed countries, is that standard 'preference-based' methods in economics typically seek to elicit individual preferences between health care options as a whole, including uncertainty and changing outcomes over time. Individual preferences involving time and uncertainty are particularly vulnerable to psychological biases and effects and can thus appear to be highly inconsistent [27]. This may be because such preferences are 'constructed' on the spot in response to context-specific stimuli [28]. Preference-based methods that ask individuals to value health care treatment options as a whole thus run a serious risk of bias [29].

The QALY approach – imposition of an additive structure of value

In the medical literature, the standard survey-based generic multi-attribute health state valuation methods used to generate the 'quality adjustment' part of QALYs – such as EQ-5D and HUI – are often referred to as 'utility' or 'preference' measures of quality of life (www.euroqol.org; www.fhs.mcmaster.ca/hug).

This helps to distinguish them from 'clinical' or 'disease-specific' quality of life measures that cannot be used to generate QALYs. This terminology also helps to emphasise that the standard methods draw on survey evidence about the opinions and values expressed by individuals (whether patients or the general public), rather than 'expert' opinion.

However, QALY data generated using a standard instrument need not be interpreted as an index of 'utility' in the standard utilitarian senses of desire-fulfilment, happiness or choice. Instead, my proposal is that QALY data generated in the standard way may be re-interpreted as representing the value of an individual's capability set – where 'value' does not necessarily have to be interpreted in terms of any of the standard utilitarian senses of 'utility'.

More specifically, I propose that the QALY can be interpreted as an index of the value of the individual's capability set in a given time period under certainty. A key feature of the QALY approach is that it imposes a high degree of structure on valuations over time and across uncertain states of nature. An individual's overall QALY gain is evaluated simply by taking the expectation of QALY gains across each uncertain state of nature and then adding up those expected QALY gains across each time period (possibly discounted). This contrasts with standard willingness to pay methods that typically allow the individual to value health care options as a whole – including all sources of uncertainty and variation over time – without necessarily imposing any separability assumptions. The QALY can thus be thought of as a 'structured' approach to valuation [30].

The advantage of a structured approach is that it helps to iron out some of the more serious biases that influence people's judgements involving small probabilities and long time periods. The problems of adaptive and constructed preferences cannot be avoided entirely, since preference data – broadly understood – are still required to provide evidence of the value of the capability set in a given time period under certainty. It would not be sensible to rely exclusively on 'expert opinion' as evidence for the value of capability sets, since expert opinions may diverge substantially from those of patients and the wider public. And, in any case, 'experts' are not immune to psychological biases.

All survey methods are vulnerable to psychological biases and effects. For example, it is well

known that different health state value elicitation techniques yield systematically different results [31]. Furthermore, standard health state valuation methods specify the health state, but not the wider capability set. So use of standard health state value elicitation methods to value capability sets means relying on the individual's own perceptions of his (unspecified) capability set and changes therein. There is no mechanism for making sure that different individuals are considering the same list of functionings when formulating their answers. Nor is there any mechanism for correcting potential misperceptions and errors in people's predictions about how changes in health status will influence their broader capability set.

One way partly to address such problems is a 'multi-method' approach [32]. Data can be obtained using a variety of standard health state value elicitation techniques – such as Visual Analogue Scale (VAS), standard gamble (SG), time-trade-off (TTO) and person trade-off (PTO). It may be possible to identify a set of 'core values' by adjusting for biases and extraneous factors such as attitudes towards risk, time and interpersonal distribution. For the purposes of the 'capability QALY', the set of core values can be thought of as representing the value of the capability set induced by specified health states in a given time period under certainty. Another complementary approach may be to use more 'deliberative' opinion polling, involving focus groups and the like, so that individuals arrive at more 'considered' responses to survey questions.

The main disadvantage of a structured approach like the QALY is that it may lack sensitivity to important interactions between different components of value over time and under uncertainty. It is well known that individual preferences between health care options do in fact violate the assumptions of additive separability across times and states of nature [33]. What matters for the capability approach, however, is the normative question of how far valuations based on capability sets should violate these assumptions. One plausible violation is that there may be a value to longevity *per se* within a single life – e.g. the ability to pursue a career, have a family, or engage in other long-term life projects – over and above the sum total value of life-years in good health taken in isolation. This might mean, for example, that saving the life of a young person might be more valuable than the sum total of QALYs gained. And there may be other violations. For

instance, holding constant the average level of capability, might it be better to have a pattern of an expanding capability set rather than a contracting one? Or a gradually changing capability set over time rather than a rapidly fluctuating one?

Sen himself says little about the structure of value over time and under uncertainty. In keeping with the flexible spirit of his approach, however, it seems sensible to leave open the possibility that capability evaluations depart from a simple adding up structure. The QALY approach can therefore only ever be an approximation to the capability approach – a ‘reference case’ analysis. Where plausible reasons can be given for valuations to violate the reference case assumption of separability across time or states of nature, the valuation should be changed accordingly. It may be a useful reference case, however, precisely because the structure it imposes is so simple and easy to understand.

The value of full health

The QALY imposes another important element of structure onto valuations, in relation to the ‘quality adjustment’ to individual life-years. It assumes that full health counts as one and death as zero for all individuals, irrespective of their wealth or talent or other characteristics. This contrasts with standard preference-based approaches, which do allow variations in the value of full health between individuals.

The primary justification for valuing full health at one is that it provides a clear and simple way of making interpersonal comparisons. This might be considered restrictive from a capability point of view, however, as two people can both have full health but quite different capability sets. For example, a rich individual with full health may have a larger and more valuable set of capabilities than a poor individual with the same level of health.

This restriction is sometimes justified in terms of equity. For example, it may be inequitable to give higher priority to prolonging the life of a rich individual rather than a poor one. Ideally, however, equity considerations should be kept separate from efficiency considerations and dealt with separately [34,35].

This restriction is likely to be of most practical concern in relation to health care decisions involving patient groups that differ substantially from the general population in terms of socio-economic status or other major determinants of non-health capabilities. On strict grounds of capability efficiency alone (i.e. maximising the aggregate value of individual capability sets, ignoring equity considerations), treating ‘diseases of the poor’ may do less good than treating ‘diseases of the rich’. This is simply because the rich have larger capability sets for any given level of health – and so stand to gain (or lose) more wider capabilities for any given change in life expectancy or health functioning. However, there are obvious counterbalancing considerations of equity such as equality of access to health care and/or equality of capability sets between socio-economic groups, which are extremely difficult to quantify [36]. So unless and until those counterbalancing equity considerations can themselves be quantified, it would seem odd to worry about quantifying the capability efficiency differential between rich and poor. So the restriction on valuing full health at one may be a reasonable assumption for most practical purposes of health care evaluation.

It is worth noting that, even with this restriction in place, QALY maximisation will still discriminate against diseases of the poor (and of the elderly). This is because treating individuals with relatively low life expectancy will yield relatively low gains in life-years [37]. As with any simple maximisation principle, therefore, QALY maximisation needs to be used judiciously and supplemented by consideration of wider concerns.

A final difficulty is how to define ‘full health’, and in particular how far to ‘relativise’ the definition to normal expectations given the individual’s age or other characteristics. For the purpose of constructing a capability QALY, it would seem appropriate to use an ‘absolute’ definition such as full health for a healthy adult (despite the difficulties this raises, for example, in relation to young children). This helps prevent the value of a capability set from depending on adaptive health expectations. It also facilitates comparability between valuations for treatments affecting different patient groups, since the valuation instrument can use a common set of descriptions of ‘full health’ for all individuals.

Differences between the ‘health QALY’ and the ‘capability QALY’

My proposal differs from the more usual interpretation of the QALY as an index of health. According to the usual interpretation, the QALY represents one important *component* of the individual’s wellbeing – the ‘health’ component. According to my re-interpretation, by contrast, the QALY represents *all* of the individual’s wellbeing. In principle, a ‘health QALY’ can sensibly be interpreted as one independent variable in a ‘wellbeing function’, alongside wealth and other variables that contribute towards individual wellbeing; whereas a ‘capability QALY’ cannot.

This may seem like splitting hairs. So some notation adapted from Broome [35] may help to clarify what important issues are at stake in choosing between the ‘capability QALY’ and the ‘health QALY’. Denote an interpersonally comparable index of an individual’s health with ratio scale properties as $h(h_1, h_2, \dots, h_m)$. This represents the value of health as a function of various health dimensions h_1, h_2 and so on. Assume that $h(\cdot)$ is a scalar with zero representing death (or a health state as bad as death). Denote full health by $H = h(H_1, H_2, \dots, H_m)$, where H_1, H_2 and so on represent the maximum level of health for each health dimension. We can then define an index of health, h/H . This is normalised so that full health is given a value of one, as is standard practice when constructing QALYs. This is the ‘health QALY’ – h/H .

Now denote an interpersonally comparable index of the value of an individual’s capability set with ratio scale properties by w (for ‘wellbeing freedom’). Again, assume that w is a scalar with zero representing death (or a capability set as bad as death). For now, however, we need not normalise or place any other restrictions on this scale of value.

The capability set, and its value w , will depend not only upon the health dimensions but also upon multiple non-health factors such as the individual’s wealth, education, local environment, personal networks, intelligence, determination and so on. For simplicity, let us assume that w can be written as a mathematical function of health and non-health variables. We can write $w = w(h_1, h_2, \dots, h_m, d_1, d_2, \dots, d_n)$, where h_1, h_2 and so on represent health dimensions and d_1, d_2 and so on represent non-health factors.

We can think of this value function as a reduced form equation, the result of a two-stage evaluation process. The first-stage models the relationship between input variables (i.e. health and resource variables) and the capability set; the second stage values the capability set along an interpersonally comparable ratio scale.

The capability set is a set of achievable scores for a list of z valuable functionings f_1, f_2, \dots, f_z . The capability set describes a z -dimensional functioning possibility frontier, on or within which the achieved functioning vector must lie. For the purpose of the current discussion we set aside the practical difficulties involved in specifying, estimating and valuing capability sets, and simply assume that the value function represents the solution to this formidable task.

The list of functionings may include some or all of the health dimensions, insofar as they are considered valuable functionings in themselves. For example, absence of pain may be a valuable state of being. The list will also include non-health functionings relating to wider aspects of life such as activities relating to work, leisure, family, politics, religion and so on and/or states of being relating to self-esteem, accomplishment, dignity, comfort and so on. Some of the non-health functionings may be hard to achieve without good health in certain dimensions; they are nevertheless separate functionings. The health variables thus play two roles: first, as input variables that influence the individual’s ability to achieve non-health functionings (e.g. the ability to play sport may depend on having good mobility) and second as valuable functionings in themselves. This is because health is an investment good as well as a consumption good [38].

Does it make sense to interpret h/H as an important component of w that can reasonably be used as a proxy for w in many health care circumstances? This would be a reasonable approximation if health and non-health factors were *separable* within the value function i.e. if we could write $w = w(h(h_1, h_2, \dots, h_m), d_1, d_2, \dots, d_n)$. If so, h would straightforwardly be a component of w . Even if w were a non-linear function of h/H , we could still try to estimate that function in order to translate changes in h/H into changes in w .

Unfortunately, however, health dimensions and non-health factors may not be separable within the capability value function. For example, the effect of decreased mobility (say, h_1) on the value of an individual’s capability to pursue valuable activities

away from home may depend on non-health factors such as wealth (e.g. ability to hire taxis), personal networks (e.g. lifts from friends) and the local environment (e.g. social service provision, physical access to amenities). By contrast, the effect of decreased pain (say h_2) on the value of capability may be independent of those non-health resources, or may depend on non-health resources in a different way. Under these circumstances, the instrumental value of health as an input variable into the capability set depends crucially on the levels of the resource variables and on which dimension(s) of health are changed. If so, an index of health *per se* can only be related to a small part of the capability set – that part concerned with the ‘intrinsic’ value of being in a good state of health. So the index of health *per se*, h/H , may not be straightforwardly related to the index of w .

An alternative is to estimate w directly, without insisting on separability between health and non-health components of value. Capability sets could be valued as a whole, taking into account both health and non-health factors. In order to be used within the QALY framework, a normalisation is required so that one represents the value of the capability set brought about by full health. We can denote the value of this capability set by $W = w(H_1, H_2, \dots, H_m, D_1, D_2, \dots, D_n)$ where D_1, D_2 and so on represent the level of non-health factors implied by full health. This takes into account the indirect effects that health variables may have on non-health factors. For instance, poor health may limit the individual’s ability to work and thus have knock-on effects on other non-health factors.

The capability QALY can then be defined as w/W – an index of the value of capability in a given time period under certainty, normalised so that one represents the value of the individual’s capability set at full health and zero represents death. In principle, w/W can be negative – if the capability set only contains functionings that are worse than death. Notice also that in principle w/W can be raised above one. This can happen if there are *direct* changes in the individual’s non-health factors or non-health capabilities that have nothing to do with changes in health. For example, if the individual receives a large cash windfall, as well as being in full health, w could rise above W . This is because W only takes into account the *indirect* effects of changes in health on the individual’s existing non-health factors and non-

health capabilities. So w/W is anchored at one, but not bounded by one.

In practice, standard generic health state valuation instruments may be closer to valuing w/W than h/H . This is because they employ broad descriptions of health dimensions that either explicitly or implicitly relate to a wide range of non-health capabilities. For example, the EQ-5D has a domain entitled ‘usual activities’ and the HUI has one entitled ‘emotion’ that explores how ‘happy’ the individual is feeling (www.euroqol.org; www.fhs.mcmaster.ca/hug). These questions cannot be answered without giving thought to non-health functionings. Furthermore, when answering questions about more narrowly health-focused domains – such as ‘pain’, ‘self-care’, and ‘visual acuity’ – respondents may naturally consider the wider impact of those domains on non-health functionings such as employment and personal relations, as well as their ‘intrinsic’ value.

The ‘capability QALY’ has another advantage over the ‘health QALY’: it can account for the direct influences that health care may have on people’s non-health-related capabilities. For example, the quality and responsiveness of personal care, social care and hotel services offered by a health care provider may influence a patient’s ability to achieve dignity, reassurance and comfort during the process of care. These influences may be quite direct, and not mediated (or at least not exclusively mediated) through their impact on health outcomes. These direct influences are ignored by h/H , which pays no attention to non-health functionings.

At least in theory, however, these direct influences can be accommodated by w/W , since w is a function of the entire capability set – including both health and non-health functionings. As noted above, direct effects on non-health functionings (i.e. ones not caused by changes in health status) may even impact so strongly as to raise w/W above one. Empirically, there are broadly two ways of modifying standard multi-attribute health state valuation instruments to incorporate these direct effects. First, explicitly to build the relevant non-health functionings (e.g. comfort, dignity) into the health state descriptions. Second, to design separate value elicitation questions that establish trade-offs between the health functionings and the relevant non-health functionings. Both approaches face serious empirical difficulties, although perhaps not insurmountable ones. Until such methods have been developed, however, standard QALY

instruments may be most useful in contexts where the health determinants of capability are of primary concern to decision-makers – such as health technology assessments that focus on a particular health care technology without varying the wider infrastructure that provides personal care, social care and hotel services.

A third advantage of the ‘capability QALY’ relates to subgroup diversity in the value of health capabilities. Individuals in different patient groups may value the same health state differently, due to differences in external characteristics (e.g. wealth, personal networks, public amenities) and internal characteristics (e.g. intelligence, willpower). An example might be the academic versus the footballer, who may place different values on mobility. The health QALY can in principle allow for such differences simply by indexing the health value function, $h(h)$, according to the individual. However, it stretches language somewhat to say that two individuals have the same ‘health state’ but different ‘health’. The capability QALY can account for such differences without such linguistic contortions, simply by noting that the value of a capability set, w/W , depends not only on the individual’s health state, h , but also on the non-health characteristics of the individual, d . Empirically, this could be achieved using sub-group analysis of survey data that relates people’s health state valuations to their non-health resources and characteristics.

Conclusion

Sen’s capability approach is a philosophically sophisticated foundation for economic evaluation that accommodates long-standing concerns with utilitarian welfare economic theory [11,39,40]. Unfortunately, it is not possible directly to apply the capability approach to economic evaluation of health care programmes. It is possible to apply the capability approach indirectly, however, by re-interpreting the QALY as an imperfect but workable index of the value of the individual’s capability set – the ‘capability QALY’.

This can sensibly be regarded as one way – although perhaps not the only way – of applying the capability approach to health care evaluation, for two reasons. First and foremost, the QALY approach pays close attention to the individual’s ‘functionings’ – what the individual can do or be.

The standard multi-attribute generic health state valuation instruments used to construct QALYs, such as EQ-5D and HUI, pay close attention to a broad set of health functionings and make implicit and sometimes explicit reference to non-health functionings. Second, the QALY approach leaves room for broader conceptions of value than the traditional utilitarian ones of choice, desire-fulfilment and happiness.

My proposal has numerous shortcomings, since the link between QALYs and capabilities is not straightforward. One shortcoming is that, like the application of Sen’s approach to the UN Human Development Index, it restricts attention to ‘elementary evaluation’. That is, it focuses on achieved functionings rather than attempting to estimate and value capability sets – the set of functionings the individual can choose from. Another shortcoming is that, in order to facilitate interpersonal comparisons, strong assumptions are made about the additive separability of value across time and states of nature, and about the value of full health.

Despite these shortcomings, the ‘capability QALY’ has advantages over the more usual interpretation of the QALY as an index of health. It acknowledges the non-separability of health and non-health components of wellbeing. It is also more flexible. It provides principled grounds for modifying or supplementing the standard methods for generating the ‘quality adjustment’ part of the QALY to account for (a) process attributes of care and other direct effects on non-health functionings, and (b) diversity in the value of the same health state to different population subgroups.

Acknowledgements

I am particularly grateful to Paul Anand, Steve Birch, Shepley Orr, Richard Smith and the anonymous referees for their detailed and helpful comments; for other helpful comments I would also like to thank John Broome, Tony Culyer, Paul Dolan, Alastair McGuire, Miranda Mugford, Adam Oliver, Mozaffar Qizilbash, Alan Williams and participants at seminars at the London School of Economics, the University of Sheffield, the University of East Anglia and the third Conference on the Capability Approach in Pavia, Italy, 7–9 September 2003 where earlier versions of this paper were presented. The usual caveat applies with particular force in relation to this paper: my errors and opinions are my own.

References

1. Sen A. Utilitarianism and welfarism. *J Philos* 1979; **LXXVI**(9): 463–488.
2. Sen A. *Commodities and Capabilities*. North-Holland: Amsterdam, 1985 (reprinted in 1999 by Oxford University Press, New Delhi).
3. Sen A. Capability and well-being. In *The Quality of Life*, Nussbaum M, Sen A (eds). Clarendon Press: Oxford, 1993; 30–54.
4. Sen A. *Development as Freedom*. Knopf: New York, 1999.
5. Sugden R. Welfare, resources, and capabilities – a review of inequality reexamined by Amartya Sen. *J Econ Lit* 1993; **31**(4): 1947–1962.
6. Robeyns I. *Valuing Freedoms: Sen's Capability Approach and Poverty Reduction*, Sabina Alkire. Oxford University Press: Oxford, 2002 (Book Review). *Econ Philos* 2003; **19**(1): 371–377.
7. Samuelson A. *Foundations of Economic Analysis*. Harvard University Press: Cambridge, MA, 1947.
8. Torrance GW, Furlong W, Feeny D. Health utility estimation. *Expert Rev Pharmacoeconomics Outcomes Res* 2002; **2**(2): 99–108.
9. Culyer AJ. The normative economics of health care finance and provision. *Oxford Rev Econ Pol* 1989; **5**(1): 34–58.
10. Culyer A. The normative economics of health care finance and provision. In *Providing Health Care: The Economics of Alternative Systems of Finance and Delivery*, McGuire A, Fenn P, Mayhew K (eds). Oxford University Press: Oxford, 1991; 65–98.
11. Sen A. *Inequality Reexamined*. Harvard University Press: Cambridge, 1992.
12. Fukuda Parr S. The human development paradigm: operationalising Sen's ideas on capabilities. *Fem Econ* 2003; **9**(2–3): 301–317.
13. Hurley J. An overview of the normative economics of the health sector. In *Handbook of Health Economics*, vol. 1A, Newhouse J, Culyer A (eds). Elsevier: Amsterdam, 2000; 55–118.
14. McGuire A. Theoretical concepts in the economic evaluation of health care. In *Economic Evaluation: Merging Theory with Practice*, Drummond M, McGuire A (eds). Oxford University Press: Oxford, 2001; 1–21.
15. Birch S, Donaldson C. Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? *Soc Sci Med* 2003; **56**: 1121–1133.
16. Boadway R, Bruce N. *Welfare Economics*. Basil Blackwell: Oxford, 1984.
17. Varian H. *Microeconomic Analysis*. Norton and Company: New York, London, 1992.
18. Griffin J. *Wellbeing: Its Meaning, Measurement and Moral Importance*. Oxford University Press: Oxford, 1986.
19. Berlin I. Two concepts of liberty. In *Four Essays on Liberty*, Berlin I (ed.). Oxford University Press: Oxford, 1969.
20. Sugden R. The metric of opportunity. *Econ Philos* 1998; **14**: 307–337.
21. Sen A. Rational fools: a critique of the behavioral foundations of economic theory. *Philos Public Aff* 1997; **6**: 317–344.
22. Williams A. Cost benefit analysis: bastard science and/or insidious poison in the body politick? *J Public Econ* 1972; **1**(2): 199–225.
23. Dolan P. The measurement of health-related quality of life. In *Handbook of Health Economics*, vol. 1A, Newhouse J, Culyer AJ (eds). Elsevier: Amsterdam, 2000; 1723–1760.
24. Burchardt T, Le Grand J. *Constraint and opportunity: identifying voluntary non-employment*. Centre for Analysis of Social Exclusion, Paper 55, London School of Economics, 2002.
25. Blackorby C, Donaldson D. The case against the use of the sum of compensating variations in cost-benefit analysis. *Can J Econ* 1990; **23**(3): 471–494.
26. Sen A. Health: perception versus observation. *Br Med J* 2002; **324**: 860–861.
27. Bell DE, Raiffa H, Tversky A. *Decision Making: Descriptive, Normative and Prescriptive Interactions*. Cambridge University Press: Cambridge, 1988.
28. Slovic P. The construction of preferences. *Am Psychol* 1995; **50**(5): 364–371.
29. Cookson R. Willingness to pay methods in health care: a sceptical view. *Health Econ* 2003; **12**(11): 891–894.
30. Broome J. Structured and unstructured valuation. *Anal Kritik* 1994; **16**: 121–132 (reprinted in Broome J. *Ethics Out of Economics*. Cambridge University Press: Cambridge, 1999).
31. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. *J Health Econ* 1996; **15**(2): 209–231.
32. Salomon JA, Murray CJL. A multi-method approach to measuring health-state valuations. *Health Econ* 2004; **13**: 281–290.
33. Loomes G, McKenzie L. The use of QALYs in health care decision making. *Soc Sci Med* 1989; **28**: 299–308.
34. Broome J. Qalys. *J Public Econ* 1993; **50**: 149–163 (also reprinted in Broome J. *Ethics Out of Economics*. Cambridge University Press: Cambridge, 1999).
35. Broome J. Measuring the burden of disease by aggregating well-being. In *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*, Murray CJL, Salomon JA, Mathers CD, Lopez AD (eds). World Health Organisation: Geneva, 2003; 89–111.
36. Sassi F, Archard L, Le Grand J. Equity and the economic evaluation of health care. In *Health Technology Assessment*, HTA Monograph Series, vol. **5**(3), 2001.

37. Anand P. The integration of claims to health-care: a programming approach. *J Health Econ* 2003; **22**: 731–745.
38. Folland S, Goodman AC, Stano M. *The Economics of Health and Health Care* (4th edn). Prentice-Hall: New Jersey, 2003.
39. Goodin R. *Utilitarianism as a Public Philosophy*. Cambridge University Press: Cambridge, 1995.
40. Hausman DM, McPherson MS. *Economic Analysis and Moral Philosophy*. Cambridge University Press: Cambridge, 1996.