Incorporating health inequality concerns into cost-effectiveness analysis - Overview

DRAFT, Revised 24 February 2012

This draft discussion paper describes a proposed analytical framework for extending cost effectiveness analysis of national health sector programmes to incorporate concerns about health inequality. The aim of this document is to describe our work in progress on this topic for discussion at a workshop to be held on 12th March 2012 in York.

An accompanying document titled 'Univariate assessment' is a more detailed technical report describing the work undertaken so far in exploring how approaches to the assessment of income inequality and poverty might be applied to health.

Project title: Identifying appropriate methods to incorporate concerns about health inequalities into economic evaluations of health care programmes

Investigators: Susan Griffin, Richard Cookson, Nigel Rice, Miqdad Asaria, Karl Claxton, Tony Culyer, Mark Sculpher, Centre for Health Economics, University of York

Funding: This project is funded through the Public Health Research Consortium under the Department of Health Policy Research Programme

1. Introduction

We intend to develop an economic evaluation framework that will enable analysts to provide quantitative information about health inequality impacts that can be used to support public sector decision making. The aim is to reflect decision makers' and stakeholders' concerns about health inequality within the context of a deliberative decision making process. It is not the aim to specify or embody any particular set of social value judgements about how health inequality reduction objectives should be formulated and traded off against other policy objectives, but rather to help analysts systematically explore the implications of alternative social value judgements.

We have chosen to focus on decision making by national health sector organisations who may be concerned about health inequality within the general population. The desirable characteristics of such organisations is that they are responsible for allocating a finite set of resources to a range of medical technologies and public health programmes, and are already familiar with the use of cost effectiveness analysis. Consequently the focus of evaluations to inform such organisations is on health benefits and health opportunity costs falling within the health sector. In the UK, for example, such organisations include the National Institute for Health and Clinical Excellence (NICE), the UK National Screening Committee, Public Health England, and the Department of Health. The benefit of assuming that there are no important non health benefits or opportunity costs falling outside the health sector is that, initially, we can set aside the additional methodological challenges of evaluating policies with inter-sectoral effects. Possible areas for case studies to illustrate our framework could include (i) screening programmes for colorectal cancer, (ii) smoking cessation programmes and (iii) blood pressure and cholesterol control programmes.

Our proposed analytical framework is divided into two conceptually distinct stages of (1) modelling health distributions and (2) ranking them. The modelling stage is about appropriately characterising the relevant health distributions, and estimating the impact of alternative interventions on those health distributions. Although modelling is partly a "scientific" endeavour, it also involves important social value judgements about the appropriate formulation of the decision options and appropriate ways of characterising the health distribution given the decision context and data availability. Appropriately characterising the health distribution involves a large number of value-laden methodological choices about social concepts and health concepts, as described below. Ranking health distributions then involves social value judgements about how to compare (appropriately characterised) distributions of health using assessment tools such as dominance rules and inequality indices.

Section 2 describes the social value judgements relevant to our proposed framework. Section 3 then describes the basic framework for cost-effectiveness analysis of health interventions that we wish to extend. Section 4 then considers some of the data requirements and analytical steps that could be used to quantify the distribution of health associated with alternative policy options.

2. Value judgements about health inequality

This project aims to provide an analytical framework that helps analysts to think systematically about the methodological choices required in order to analyse health inequality. A crucial aspect of this involves being as clear and explicit as possible about the contestable value judgements that are embodied in alternative methodological choices. However, we do not aim to provide guidance on appropriate deliberative processes through which decision makers and stakeholders might arrive at those value judgements and communicate them to analysts.

2.1 General issues

Decision makers' concerns about health inequality reflect value judgements about how far different kinds of health inequality are fair, or just or equitable. (We take these three terms – fairness, justice and equity – as approximate synonyms.) There is no prospect of agreement on a fully specified societal set of such value judgements for two reasons: i) reasonable people disagree about what counts as "justice" so there is no consensus on a complete specification of distributions that could be regarded as just; and ii) how far any particular health inequality is considered unjust often depends on contestable factual beliefs about the causes or determinants of that health inequality.

The causal pathways leading to health inequality are often complex, making it difficult to gather conclusive evidence on cause and effect despite conclusive evidence of association. For example, there is overwhelming evidence of a positive association between income and health within all countries – one of the many ubiquitous and persistent "social gradients in health" that epidemiologists and economists have documented. Yet the causal pathways that generate that association remain disputed. Low income may cause poor health through cumulative, multi-component and multi-stage pathways, involving effects of unhealthy material living and working conditions and psycho-social stress due to low position in the social hierarchy, all of which may, in turn, cause poor diet, limited physical activity, smoking and harmful drinking, and other forms of unhealthy behaviour. There may also be "reverse" causal pathways: poor health may cause low income due to reduced ability to work and develop a well paid career. Health and income may also be jointly determined: third factors may cause both poor health and low income – such as a damaged childhood, poor education, low innate ability, or a high rate of time preference leading individuals to choose low rates of investment in both human and health capital.

When analysing health inequality concerns, therefore, analysts make methodological choices which implicitly or explicitly embody highly contestable scientific and social value judgements. It is therefore incumbent on analysts to ensure that (A) the value judgements embedded in their analyses are clear and explicit, and (B) a range of scenarios are provided which reflect alternative but reasonable views about justice and the nature of the underlying causal relationships. Presenting

analysis in this way can serve as a useful starting point for deliberation, and could be used to provide a record of the analysis and corresponding reasoning that lay behind the final decision, which could facilitate transparency and consistency with previous and future decisions

We can distinguish at least three different kinds of value-laden choices that must be made in order to analyse health inequality concerns:

1. ***"Inequality between whom?"*** The selection of **social concept(s)** to characterise unjust health inequality (e.g. inequality in health associated with differences in income, education, area deprivation or some other social determinant of health) and a description of any accompanying beliefs about causal pathways that underpin that selection.
2. ***"Equality of what?"*** The selection of **health concepts(s)** (e.g. whether to focus on the distribution of health gains and losses or health levels, cause-specific health or all-cause health, current health or lifetime health).
3. ***"Equality measured how?"*** The selection of **inequality concept(s)** (e.g. whether to focus on improving the health of the worst off or narrowing the distribution; e.g. whether to focus on relative inequality, absolute inequality or shortfall inequality concepts; and so on).

In practice, such choices may be driven or constrained by data limitations. For example, one might want to know how adult health varies by household income in childhood but in practice only have data on how adult health varies by area deprivation score in adulthood. Nevertheless, it is important to be clear about the value judgements implied by whatever imperfect data and methods are actually used in practice.

We can distinguish three broad frameworks for health inequality measurement, within which these value-laden methodological choices are embedded:

1. **Univariate analysis** (Gakidou, Murray, & Frenk, 2000). This approach ignores the determinants of health and instead focuses exclusively on the resulting distribution of health, ideally observed or estimated for all individuals in the general population – for example, the distribution of age at death or health expectancy. This describes pure health inequality.
2. **Bivariate analysis** (Wagstaff, Paci, & van Doorslaer, 1991)**.** This focuses on the joint distribution of health together with another variable – such as income – which decision makers consider to be a relevant source of unjust health inequality. The second variable may be continuous or categorical; the unit of analysis may be individuals or groups; and the joint distribution may or may not be standardised to allow for health determinants such as age that are not considered to be unfair sources of health inequality.
3. **Multivariate or "standardised univariate" analysis** (Fleurbaey & Schokkaert, 2009). This first standardises the distribution of health with the aim of purging multiple fair causes of

health inequality while preserving multiple unfair causes of health inequality. It then focuses on the standardised distribution of health, to which it becomes possible to apply univariate measures of inequality.

In practice, there may be little support for the rather extreme view that all variation in health is equally unjust and the cause of health variation does not matter, and hence the relevance of 'pure' univariate analysis may be limited (Braveman, Krieger, & Lynch, 2000). For this reason, proponents of univariate analysis often argue that it should be tempered by a form of standardisation and/or supplemented by bivariate analysis. For example, (Gakidou et al., 2000) argue for a focus on health expectancy rather than (quality-adjusted) age at death, precisely because the former is purged of variation due to unavoidable random accidents.

Bivariate analysis remains a common approach to health inequality measurement in both the economic and epidemiological literatures, and numerous different bivariate indices of health inequality have been proposed and used. However, (Fleurbaey & Schokkaert, 2009) argue that standardised univariate analysis has at least three advantages over bivariate analysis:

- It encompasses and allows simultaneously for multiple sources of unjust health inequality, avoiding excessive focus on one particular source
- In order to undertake standardisation analysts must specify scientific value judgements about the causal pathways between social variables and health, and also social value judgements about the injustice of those causal pathways
- Social value judgements about which determinants of health count as unjust (i.e. made in the standardisation process) can be separated from social value judgements about the magnitude of inequality (i.e. made in assessing inequality in the resulting distribution).

We aim to illustrate the similarities and differences between the three approaches, showing when they will produce similar rankings of alternative policy options, and explaining the differences and interpretation when they do not.

2.2 Inequality between whom

It is important to consider three main issues when assessing how far a particular social determinant of health may be an unjust source of health inequality:

1. How far the resulting health inequality is **remediable** (or "avoidable" or "preventable" or "amenable" or "socially controllable")
2. How far the resulting health inequality is **compensable** by non-health benefits such as income
3. How far the resulting health inequality is a matter of **individual responsibility** (or "choice" or "effort")

The philosopher Daniel Hausman has emphasised the importance of the first two issues, arguing that "health inequalities that are neither remediable nor compensable – for example, inequalities due to conditions such as Tay Sachs disease – are tragic but not unjust" (Hausman, 2007). Although decision makers may not take such a strong view, they will certainly want to consider how far they are capable of doing anything about the health inequality in question – either by reducing the health inequality, or by providing some form of compensation to those who suffer from the health inequality.

Factors that affect how far any given health inequality is remediable or compensable by a particular decision maker include:

- The availability of effective possible actions to remedy the health inequality
- Political constraints on what possible actions for remedying health inequality are feasible
- The time horizon – for instance, some health inequalities may not be remediable within a few years but may be remediable over decades or generations
- The opportunity costs of remedying or compensating the inequality
- The remit of the relevant decision making body – for instance, health sector decision makers have little or no scope for altering decisions in other sectors that influence health inequality such as taxation or education
- The scope of the current decision problem – for instance, the health inequality may not be remediable by any of the available options relating to the particular decision in hand, even though future decisions on other topics within the broad remit of the decision making body could potentially alter the health inequality

To the extent that a decision maker faced with a particular decision problem is not capable of remedying or compensating a particular source of health inequality, there may be little point in incorporating that source of health inequality into economic evaluations designed to inform that particular decision. Once deemed capable of remedying a particular source of health inequality there may be further questions as to whether a particular decision maker ought to. This would require consideration of the normative frameworks that underlie the role of the decision maker in terms of remit and objective function. Decision makers will also of course be interested in the opportunity costs of remedying or compensating the inequality – i.e. what sacrifices may have to be made in terms of other policy goals – and they may be interested how far other decision making bodies are capable of remedying or compensating the health inequality more effectively or at a lower opportunity cost. These issues will be particularly relevant when considering how to extend our proposed framework to reflect intersectoral comparisons. Some philosophers argue that assessment of health justice "must precede considering what our current social institutions are capable of addressing" (Venkatapuram, 2011). It may indeed sometimes be useful to adopt a "God's eye" perspective, unconstrained by the political, economic and technological constraints faced by current decision makers and free to

envision the shape of a future, better society. However, economic evaluations designed to inform current actual decision making must also adopt the perspective of current actual decision makers.

The third issue that decision makers may want to consider is how far the health determinant is a matter of individual responsibility. The "liberal-egalitarian" tradition in both philosophy and economics emphasises the importance of assessing equality of opportunity for outcomes rather than merely equality of outcomes. The central issue in this literature is distinguishing how far outcomes are due to circumstances for which individuals ought not to be held responsible versus choices or effort for which individuals ought to be held responsible. Insofar as decision makers believe that sources of health inequality that are entirely a matter of individual responsibility are not unfair, it may be appropriate to purge this source of variation from the health inequality analysis. Economists have started to develop methods for doing this empirically, in attempting to disentangle multiple "circumstance" variables (such as parental social class and country of birth) from multiple "effort" variables (such as smoking and harmful drinking). This work inevitably involves contestable value judgements, of course, since "effort" variables may be partly caused by "circumstance" variables. Researchers have started to develop empirical methods for addressing this thorny difficulty and for placing upper and lower bounds on the extent of overall equality of opportunity (Rosa Dias, 2009). It is fair to say, however, that such empirical work is embryonic and still some way from offering directly applicable parameter estimates for use in economic evaluation studies.

2.3 Equality of what

Our second set of value laden choices involves the selection of health concepts for describing health distributions. These require consideration of:

- The measurement scale of the basic unit of health – for instance, ordinal or a fixed ratio scale
- The nature of the basic unit of health - for instance, mortality or quality-adjusted life years
- The time period for measuring the basic health unit - for instance, the current year, the future remaining lifetime from the point of decision, or the whole lifetime including past, current and expected future health
- The scope of the basic health unit - for instance, whether to look at health related to bowel cancer only, or all cancers, or all causes of ill health.
- The form taken by the basic health unit in the final health concept of interest from a distributional perspective, in other words, the "distribuendum" - for instance, changes in health or achieved levels of lifetime health or shortfalls from a target level of lifetime health.

Different ways of defining the distribuendum imply different social value judgements about justice in the distribution of health and will result in different scientific conclusions about the extent of inequality. For example, imagine we are interested in quality-adjusted life years, measured over the

whole lifetime, and related to all causes of ill health. If one were to focus on the distribution of gains and losses in lifetime quality-adjusted life years brought about by an intervention this will not necessarily indicate where resources ought to go to reduce inequality in lifetime health levels.

Standard cost effectiveness analysis requires that all outcomes be expressed in terms of a common numeraire. Correspondingly, we focus on health measured on a cardinal scale. Sometimes, however, health distributions may be described using a health variable measured on an ordinal scale, such as self-reported health. If so, it may still be possible to analyse health impacts, but it may be difficult to specify the nature of trade-offs between the improvement in health offered by an intervention and the opportunity cost in terms of displaced activities, both in terms of overall health gains and in terms of any inequality reduction.

2.4 Equality measured how

Further value judgements are embodied within the choice of measure of inequality. Broadly, measures of inequality can be divided into those that describe relative inequality (e.g. a ratio between two groups), those that describe absolute inequality (e.g. a gap or difference between two groups), those that describe shortfall inequality or poverty (e.g. the number of people who fall short of a target level of health), and those that are simultaneously sensitive to more than one of these three dimensions of inequality. A given change in a distribution of health can lead to an increase in relative inequality and a decrease in absolute inequality, or vice versa. Therefore, in order to say whether inequality is improved it is necessary to specify the nature of the inequality concern in terms of whether it is relative or absolute or shortfall inequality aversion or a combination. This will determine the appropriate distribution to report. The impact of an intervention on absolute inequality can be determined from the distribution of the change in health associated with that intervention or from the distribution of total health levels. However, the impact of an intervention on relative or shortfall inequality requires additional information on levels of health.

It may be possible to define some dominance criteria which may avoid having to select a particular inequality concept. For example, decision makers may judge that the extent of any inequality in the resulting distribution of health is not important were an intervention found to be "Pareto dominant" in the sense of being better for the health of all individuals in the general population. Alternatively it may be that some interventions would not only increase overall population health but also reduce health inequality of all kinds – a 'win-win' situation. A decision rule to reimburse such interventions would be consistent with a broad range of underlying social welfare functions that are increasing in health and decreasing with respect to inequality in health. However, this would not produce a complete ranking of all interventions. Additional value judgements would be required in order to choose between an intervention that is not cost effective (i.e. reduces overall population health) but reduces inequality (of one or more kinds) and one that is cost effective (i.e. increases population

health) but also increases inequality. These additional value judgements begin to describe how much improvements in equality are preferred in terms of health forgone. It will be necessary to consider whether and how it may be possible to compare health inequality reduction between studies that use different social variables, different health metrics and different inequality measures.

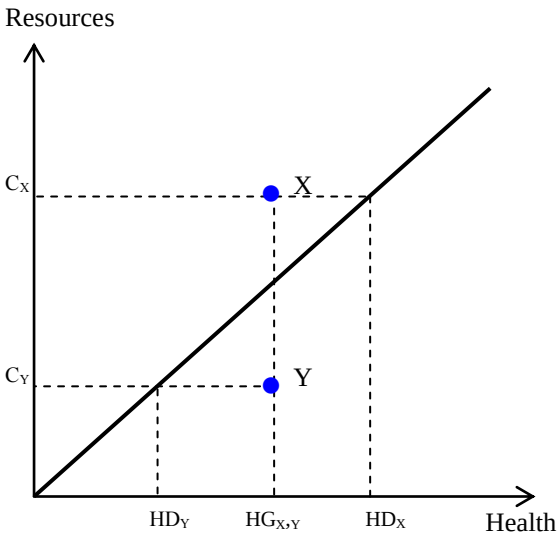
3. Analytical framework for standard cost-effectiveness analysis

The role of cost-effectiveness analysis is to provide information to decision makers to help them identify which of a set of investment opportunities would provide the greatest payoff. The context of this research is to provide an analytical framework for cost-effectiveness analyses that aims to inform decision makers charged with allocating resources to health care technologies and public health programmes (which we shall refer to collectively as health interventions). We take as a starting point a societal decision making framework and describe a typical cost-effectiveness analysis conducted from a health sector perspective. That is, we assume a fixed budget for the provision of health interventions, and that the interventions do not have important non-health benefits or impose costs outside the health sector. One objective of the decision maker is assumed to be maximisation of expected population health levels from available resources. Another, we assume, is to make the distribution of health more equal. Total population health is regularly calculated as the sum total of expected individual discounted health levels. This typifies the use of cost-effectiveness analysis by health technology assessment agencies such as NICE. Taking the methods used by NICE as a starting point, we describe in more detail how cost-effectiveness analysis is currently undertaken when informing decisions aimed at health maximisation before considering extensions to incorporate health inequality concerns. It is important throughout to consider whether there are elements of this initial analytical framework that cannot be maintained and thus require revision rather than extension when the objective is no longer simply to maximise population health levels.

Theoretically, cost effectiveness analysis could take the form of estimating, for all possible interventions for all potential recipients, the expected health outcomes and cost burden and comparing this to the total amount of resources available. The whole allocation problem could then be solved simultaneously by using mathematical programming techniques. However, the informational requirements of this 'global', system-wide approach make it impractical given the need to inform decisions in a timely manner. A more practical 'case-by-case' approach is to consider the choice between the interventions that are relevant to a more narrowly defined decision problem, for example the treatment options available for a specific indication or disease. Even then the ideal informational requirements are rarely fully met. Should one of the mutually exclusive options be found to offer an improvement in health but at a greater cost than the currently funded alternative, the potential improvement in health outcomes must be compared to the health lost as a result of discontinuing interventions by reallocating resources to fund the new intervention.

Health outcomes are quantified in terms of a single cardinal measure of health that reflects mortality and morbidity, for example a quality adjusted life year. The purpose is to have a common measure that can be applied to all evaluations, allowing comparisons to be made across a range of health conditions. The resources of interest, which determine the cost burden of the interventions, are simply those that fall on the budget constraint, i.e. resources within the health sector. The budget is fully allocated, so introducing a new health intervention displaces another. The health gains forgone by displacing existing health interventions represent the opportunity cost of introducing the new intervention. If the health gains from the new intervention exceed this opportunity cost then it is regarded as cost-effective. Box 1 describes the process of evaluation.

Box 1. Evaluating the cost-effectiveness of new health interventions



The slope of the diagonal line represents the amount of health that can be generated per unit of resources used to provide existing health interventions. Two new interventions (X and Y) are evaluated to determine whether they should be provided within the health sector.

Intervention X offers a health gain HG_X at a cost of C_X. Reducing provision of existing services to fund X would displace amount of health HD_X, resulting in a loss of health overall as $HG_X - HD_X < 0$

Intervention Y offers a health gain HG_Y at an additional cost of C_Y. Reducing provision of existing services to fund Y would displace amount of health HD_Y, resulting in a gain of health overall as $HG_Y - HD_Y > 0$

As an example we consider a cost-effectiveness analysis of intervention X from Box 1 in which it is compared to current practice. It is estimated that intervention X increases health gains in recipients by 0.7 quality adjusted life years per patient (HG_X=0.7) at an additional cost per patient of £16,000 relative to current practice. The set of interventions currently funded in the health service are assumed to generate quality adjusted life years at a rate of one per £20,000 worth of resources invested. Using this cost-effectiveness threshold the health opportunity cost per patient receiving intervention X can be calculated as $\frac{£16,000}{£20,000} = 0.8$ quality adjusted life years (HD_X=0.8). The per patient net health benefits of intervention X are therefore $NHB_X = 0.7 - 0.8 = -0.1$. Multiplying this by the size of the target population would provide the expected overall loss in health benefits from diverting resources from currently funded interventions in order to reimburse intervention X.

Identifying and evaluating displaced health service activity is not often undertaken. Instead an estimation of the shadow price of the budget constraint provides a threshold for assessing cost-effectiveness. This threshold represents an estimate of the impact of the displacement that would occur by diverting resources to the new intervention in terms of health forgone, i.e. the opportunity cost of resources in health terms. When the amount of additional resources demanded by each intervention has only a marginal impact on existing services the cost-effectiveness threshold will not alter with each reallocation of resources, and a common estimate can be used within each budgetary period. Thus a common cost-effectiveness threshold can be used to describe the cost burden of interventions in terms of health, and these can be subtracted from the expected health gains attributed to an intervention in order to describe its net health benefit. This approach means that cost-effectiveness analyses identify and characterise the eligible patient population for the intervention being evaluated, but the characteristics and health levels of non-recipients are left undetermined.

Given an objective to maximise population health gains from available resources, a decision rule to reimburse the interventions with greatest expected net health benefits would achieve this objective and produce a complete, transitive ranking of alternative health interventions. The optimal set of interventions could be identified by evaluating total expected lifetime health levels. However, the changes in health levels between the interventions being compared are more commonly calculated from the point at which the treatment decision is made, leaving health achieved up to that point unspecified. From the point of the treatment decision outcomes that do not differ according to the intervention received may also be left unspecified. So, typically interventions may be evaluated by focusing on changes in health, without explicitly characterising the resulting levels of health.

3.1 Data requirements for standard cost-effectiveness analysis

The evidence and methods of calculation must distinguish the health gains that are attributable to the use of different interventions. Estimation of treatment effects, that is the change in health outcomes or resource use that would result from the use of an intervention, requires evidence on the causal effects of interventions. Average treatment effects may also be reported for sub-groups based on clinically relevant characteristics which are known at the point of treatment decision such as gender, presence of co-morbidities, or risk factors. This is often described as treatment effect heterogeneity because the variation from the overall mean treatment effect is explained by known patient characteristics. Sub-groups can also be defined in terms of characteristics that determine baseline risk or prognosis in order to estimate absolute changes in health. In order to maximise population health levels sub-group specific net health benefits should be estimated if it is possible that different treatment decisions would be taken between groups.

Where evidence on treatment effectiveness pertains to clinical outcomes (e.g. survival, progression-free survival, or events averted), additional evidence is required to translate this into relevant health

outcomes, for example by extrapolating to a relevant time horizon and applying health related quality of life weights. The cost burden of an intervention can be described in terms of the cost of providing the intervention itself and other associated health care resources utilised over the course of a patient's lifetime. Evidence on resource use is converted into total costs by applying appropriate unit costs. Most commonly, elements of health care resource use and health outcomes that are not expected to differ between the interventions under comparison are left unspecified.

In practice the process of identifying the available evidence and the manner in which it is interpreted and used to calculate expected costs and health outcomes will be subject to uncertainty. Uncertainty in estimates of cost-effectiveness leads to the possibility that the wrong set of interventions is funded such that population health is not maximised, and hence there is an opportunity cost to uncertainty in terms of health forgone. The uncertainty in evidence can be characterised by describing the range and likelihood of alternative plausible values and assumptions supported by the evidence, and from this computing the range and likelihood of expected costs and health outcomes for each treatment under evaluation. It is then possible to address questions about whether the current evidence base is sufficient to support the choice between alternative treatments and whether further research is valuable. This uncertainty around the expected costs and health outcomes is distinct from the variation in costs and health outcomes within the population, and it is the latter that is the focus of this research. For simplicity we will consider that the distribution of outcomes between individuals or sub-groups is known (i.e. is not uncertain), but will return to this issue of uncertainty at later stages in the overall research grant.

4. Extending the framework to incorporate health inequality concerns

We now ask how this framework can be extended to incorporate inequality concerns. Additional data is required for cost-effectiveness analyses to characterise distributions of health as well as overall population health levels. As described earlier, there are a range of health distributions that could be of interest to decision makers. However, the main modelling issues can all be illustrated by focusing on two basic health distributions:

- the distribution of changes in health attributable to the intervention
- the distribution of lifetime health levels with and without the intervention

We illustrate the additional requirements for characterising these two basic health distributions using a simple numerical example, which extends the cost effectiveness analysis of intervention X from Box 1. Our discussion is organised around Table 1, which sets out the key additional pieces of information that would describe these two basic distributions.

Table 1. Extending standard cost-effectiveness analysis to evaluate distributional impacts.

		Standard cost-effectiveness analysis			Population distribution		Health after treatment age (40)		Health before 40	Total health level	
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Rank	Recipient	Intervention gain	Intervention cost	Gain-cost (A-B)	Opportunity cost	Net health benefit (A-D)	QALE baseline	QALE with intervention (F+E)	QALYs accrued	QALE baseline (F+H)	QALE with intervention (F+H+E)
1	Yes	**0.5**	1	-0.5	**0.4**	0.1	28	28.1	36	**64**	64.1
2	Yes	**0.6**	1	-0.4	**0.4**	0.2	29	29.2	36	**65**	65.2
3	Yes	**0.7**	0.8	-0.1	**0.4**	0.3	30	30.3	37	**67**	67.3
4	Yes	**0.8**	0.6	0.2	**0.4**	0.4	32	32.4	37	**69**	69.4
5	No				**0.4**	-0.4	33	32.6	38	**71**	70.6
6	No				**0.4**	-0.4	34	33.6	38	**72**	71.6
7	No				**0.4**	-0.4	35	34.6	39	**74**	73.6
8	Yes	**0.9**	0.6	0.3	**0.4**	0.5	36	36.5	39	**75**	75.5
9	No				**0.4**	-0.4	38	37.6	40	**78**	77.6
10	No				**0.4**	-0.4	40	39.6	40	**80**	79.6
Recipient avg		0.7	0.8	-0.1	0.4	0.3	31	31.3	37	68	68.3
Non-recipient avg					0.4	-0.4	36	35.6	39	75	74.6
Total population		**3.5**	**4**	**-0.5**	**4**	**-0.5**	335	334.5	380	715	714.5

Table 1 describes a total population of 10 homogenous groups of individuals who are ranked according to their total expected health level without the intervention.

Column A describes the true distribution of health gains attributable to the intervention

Column B describes the true distribution of additional resources (in health terms) invested in the intervention

Column C describes how 'net health benefits' are calculated in 'standard' cost-effectiveness analysis – i.e. implying that health opportunity costs fall on the individual patient who utilises an amount of resources. This does not happen, as the NHS does not have a fixed expenditure limit per patient

Column D describes the true distribution of the health opportunity cost of the additional resources used to provide the intervention

Columns F and G describe the distribution of health levels from the age at which patients would receive the intervention

Columns I and J describe the distribution of total health levels with and without the intervention

4.1 Distribution of changes in health

Existing cost-effectiveness analyses characterise the group of patients that is eligible to receive the interventions being evaluated. For example, a cost-effectiveness analysis of treatments for lung cancer would include patients diagnosed with lung cancer and would exclude anyone that did not develop lung cancer. We will refer to patients eligible to receive the intervention in question (at any point during their lifetime) as the 'recipients', and the remaining proportion of the population who will never be eligible for the intervention as 'non-recipients'. If the same policy applies across the whole population eligible to receive an intervention then cost-effectiveness analyses may report only the average gain and cost per recipient (i.e. the average treatment effect for the treated). A distribution of changes in health attributable to a particular intervention may in part be characterised by existing cost-effectiveness analyses that incorporate sub-group specific health benefits, or such analyses could readily be extended to do so. For example, we might split the recipients of intervention X into five separate groups and estimate health benefits and costs for each as shown in columns A and B of Table 1. A distribution of health gains and resource use within recipients of the intervention can be fully described by characterising systematic differences in treatment effect and baseline risk according to patient characteristics, regardless of whether those characteristics can be used to determine the treatment decision. In principle, this could be specified in terms of health outcomes and costs for every individual in the population. However we use a group level example here for exposition, and to recognise that in practice the availability of data may limit our ability to distinguish differences beyond the broader level of homogenous groups of individuals with similar sets of characteristics.

This distribution of health gains can then start to be placed in the context of the overall population by characterising which groups within the overall population would be expected to receive the intervention. For example, we estimate that 50% of the population would be diagnosed with the disease targeted by intervention X. We then need to know the distribution of health opportunity costs in the overall population. By using the cost-effectiveness threshold to describe the intervention costs in health terms it is possible to describe a distribution of 'net health benefits' as shown in column C. However, this distribution of net health benefits in column C does not in general describe a distribution of net changes in health attributable to the intervention. That would only be the case if any additional resources utilised by a group receiving intervention X were drawn from other health care activities that would have been received by that same group, i.e. if the total investment of health care resources in each group were fixed. This is not the case within a collectively funded health system where only the total budget is fixed and therefore the opportunity cost of the resources used to provide an intervention does not necessarily fall on the patient on whom those resources are spent.

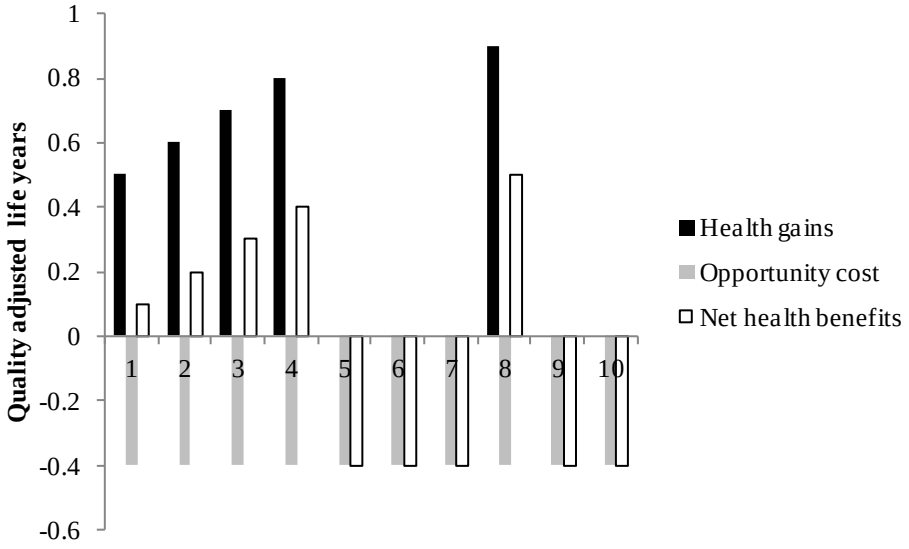
- Existing cost-effectiveness analyses may be able to describe a distribution of health gains to intervention recipients, but they can describe only the total opportunity cost and not its distribution among either recipients or non-recipients

In order to characterise the net distributional impact of a new intervention on health changes it is necessary to describe both the distribution of the health gains from the intervention and the distribution of the opportunity cost resulting from activities displaced to free up budget (assumed to be fully allocated) to carry out the intervention. The total amount of additional resources required to fund intervention X is the number of recipients multiplied by the per patient cost $5 * £16,000 = £80,000$. How this affects the distribution of health will depend on who would have benefited from existing interventions that are displaced. It could be the case that a general estimate of the distribution of health losses from displaced activities could be used across all cost-effectiveness analyses in a similar manner to the cost-effectiveness threshold. This could then be used to convert the total intervention cost into a distribution of opportunity costs. In this example the impact of displacement is assumed, simply, to be distributed evenly across the whole population as shown in column D of Table 1 (discontinuation of activities worth $£80,000$ which displaces $\frac{£80,000}{£20,000} = 4$ units of health, distributed evenly across a population of 10 groups).

- The distribution of the opportunity cost from displaced activities must be estimated

In general individuals who are not among the recipients for the interventions being evaluated would suffer a net health loss when resources are reallocated to the previously unfunded intervention. Using the distribution of health gains from the intervention and the distribution of opportunity costs it is possible to describe the distribution of net changes in health attributable to the intervention, as shown in Figure 1. This actual distribution of net health benefits is different to that which would be described by looking at the sub-group net health benefits from an 'unextended' cost-effectiveness analysis which implied that the opportunity cost fell on the same patients who utilised additional resources.

Figure 1. Distribution of change in health with intervention



4.2 Distribution of lifetime health levels

In existing cost-effectiveness analyses it is common to estimate quality adjusted life expectancy from the time at which the individual receives the intervention. So, if the cost-effectiveness analysis of intervention X considers patients aged 40 years, it may be able to describe the remaining quality adjusted life expectancy among potential recipients without the intervention (i.e. at baseline) and to break this down by sub-group as shown in rows 1-5 of column F in Table 1. To do this:

- Additional information may be required on the distribution of remaining quality adjusted life expectancy among sub-groups of the intervention recipients
- Existing cost-effectiveness analyses may capture this in sub-group analyses, but the number of sub-groups may need to be expanded to incorporate characteristics that are do not impact on incremental net health benefit

To move from changes in remaining health levels to total expected lifetime health levels it is also necessary to have information on achieved health levels up to the point at which the treatment decision is made. That is, it is necessary to know the starting distribution of health among the population, as shown in rows 1-5 of column H in Table 1. Unless this information is relevant for estimating future health gains, it is unlikely to be available from existing cost-effectiveness analyses.

- Additional information may be required to estimate any inequality in health achieved by those eligible to receive an intervention being evaluated

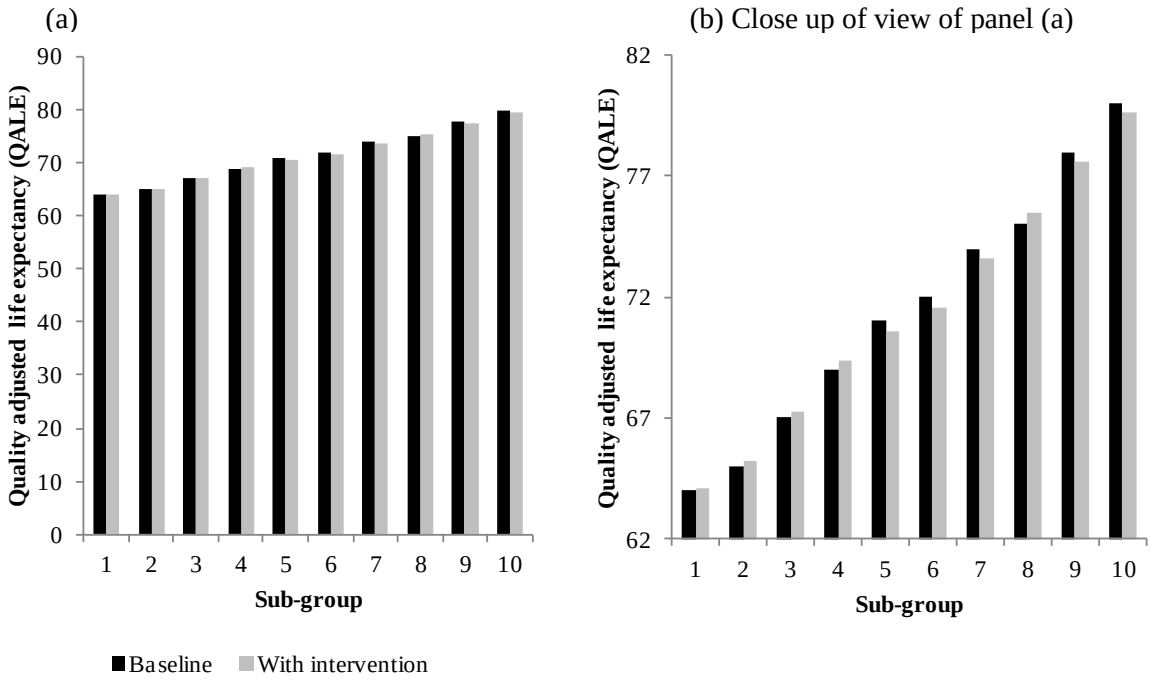
Information on health levels will also be required for the remaining population who would not receive the intervention. By incorporating additional information on quality adjusted life expectancy after age

40 among non-recipients this would complete column F in Table 1. Similarly, quality adjusted life years accrued up to age 40 could be used to complete column G for non-recipients. However, while the defining patient characteristics at the point at which the treatment decision is made has relevance when estimating health gains from the intervention, a similar approach to estimating opportunity losses for non-recipients cannot be taken unless the interventions that are displaced are identified. Instead total health levels may be estimated directly for the non-recipients without identifying the point at which the discontinued activities would have been received. In practical terms, if the distribution of changes in health attributed to the intervention has been estimated accurately it may also be convenient to estimate directly the total baseline health levels without the intervention for the whole population.

- Additional information is required on total expected health levels in the population

The distributions of total expected total health levels with and without the intervention, as shown in columns I and J of Table 1 are illustrated in Figure 2. The differences between the health levels achieved by each person with and without the intervention in Figure 2 correspond to the full population distribution of net health benefits in Figure 1. However, the scale is adjusted in order to show them in the context of health levels that do not differ according to the intervention.

Figure 2. Distribution of health with and without intervention



In summary, at a minimum an extended cost-effectiveness analysis could estimate the emboldened cells in columns A and B of Table 1for a set of mutually exclusive interventions aimed at a particular

patient population. It may then be possible to incorporate additional external information regarding the incidence of the disease and the emboldened cells in columns D and I in order to fully describe the two distributions of interest across the entire population. For example, intervention X would not be regarded as cost-effective in terms of maximising population health, as it is estimated to reduce population health by 0.5 quality adjusted life years (0.05 per person). Thus it would not be reimbursed based on a decision rule to maximise population health. The re-distributive impact of intervention X is to increase net health benefits among recipients by 0.3 quality adjusted life years (0.06 per person) and reduce net health benefits in the non-recipients by 0.4 quality-adjusted life years (0.08 per person). The weighting implied by assessment criteria under which intervention X would be reimbursed is that the value of health benefits to recipients exceeds those in the non-recipients by 4:3. Similar trade-offs could be reported based on criteria relevant to the specific inequality concern, for example based on expected total health levels if the aim is to prioritise the worst off.

The extensions to the framework here suggest that in order to describe an individual level distribution of health it is necessary to characterise any variation explained by patient characteristics in:

1. treatment effect in terms of health gains and resource use;
2. incidence or eligibility for the interventions being evaluated;
3. opportunity cost;
4. total health levels without the intervention

This would allow the extent of each of these to be estimated for each individual or sub-group within a population. A value judgement is required as to which inequalities in the distribution of health are regarded as unfair. If the inequality concern relates to any variation in health outcomes then a cost-effectiveness analysis needs to describe the distribution of health outcomes rather than the average or total expected level. However, if the inequality concern does not apply to all health, then further analysis is required in order describe the distributional impact on unfair variation in health outcomes. For example, differences in health outcomes could be presented for relevant sub-groups between whom equality in health is desired. Alternatively, this could take the form of a multivariate analysis in which estimates of total health are adjusted to control for fair differences in order to describe the unfair distribution of health. This requires that additional data be collected on equity relevant characteristics alongside data on treatment efficacy, health-related quality of life and resource use. The number of potentially relevant characteristics could be very large, and data availability may be limited. The selection of which characteristics to include in cost-effectiveness analysis and the level of detail required in the estimated distribution in health will need to be determined pragmatically and according to underlying value judgements.

5. References

Braveman, P., Krieger, N., & Lynch, J. (2000). Health inequalities and social inequalities in health. *Bulletin of the World Health Organization*, *77*(7), 232-235.

Fleurbaey, M., & Schokkaert, E. (2009). Unfair inequalities in health and health care. *Journal of health economics*, *28*(1), 73-90. doi:10.1016/j.jhealeco.2008.07.016

Gakidou, E. E., Murray, C. J., & Frenk, J. (2000). Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bulletin of the World Health Organization*, *78*(1), 42-54.

Hausman, D. (2007). What's Wrong with Health Inequalities?*. *Journal of Political Philosophy*, *15*(1), 46-66.

Rosa Dias, P. (2009). INEQUALITY OF OPPORTUNITY IN HEALTH : EVIDENCE FROM A UK COHORT STUDY. *Health Economics*, *1074*(July), 1057-1074. doi:10.1002/hec

Venkatapuram, S. (2011). *Health Justice: An Argument from the Capabilities Approach* (p. 288). John Wiley & Sons.

Wagstaff, A., Paci, P., & van Doorslaer, E. (1991). On the measurement of inequalities in health. *Social science & medicine (1982)*, *33*(5), 545-57.