THE UNIVERSITY of York

Mark versus Luke?
Appropriate Methods for the Evaluation of Public Health Interventions

CHE Research Paper 31

Mark versus Luke? Appropriate methods for the evaluation of public health interventions

KP Claxton [1, 2]
M J Sculpher [1]
AJ Culyer [2, 3]

Centre for Health Economics, University of York, UK [1]
Department of Economics and Related Studies, University of York, UK [2]
Institute for Work and Health, Toronto, Ontario, Canada [3]

November 2007

Background

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The new CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

This work was funded by a programme grant from the UK Medical Research Council as part of the Health Services Research Collaboration.

An earlier version of this paper was presented at the Health Economists Study Group meeting held in York, July 2006. We would like to acknowledge the helpful discussion and comments from colleagues. In particular we would like to thank Hugh Gravelle for insightful comments on an earlier draft. All errors of omission and commission remain our own.

Disclaimer

Papers published in the CHE Research Paper (RP) series are intended as a contribution to current research. Work and ideas reported in RPs may not always represent the final position and as such may sometimes need to be treated as work in progress. The material and views expressed in RPs are solely those of the authors and should not be interpreted as representing the collective views of CHE research staff or their research funders.

Further copies

Copies of this paper are freely available to download from the CHE website http://www.york.ac.uk/inst/che/publications/index.htm. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third-parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk, telephone 01904 321458 for further details.

Centre for Health Economics
Alcuin College
University of York
York, UK
www.york.ac.uk/inst/che

Abstract

The purpose of this paper is to demonstrate that a social decision making approach to evaluation can be generalised to interventions such as public health and national policies which have multiple objectives and impact on multiple constraints within and beyond the health sector. We demonstrate that a mathematical programming solution to this problem is possible, but the information requirements make it impractical. Instead we propose a simple compensation test for interventions with multiple and cross-sectoral effects. However, rather than compensation based on individual preferences, it can be based on the net benefits falling on different sectors. The valuation of outcomes is based on the shadow prices of the existing budget constraints, which are implicit in existing public expenditure and its allocation across different sectors. A 'welfarist' societal perspective is not sufficient; rather, a multiple perspective evaluation which accounts for costs and effects falling on each sector is required.

Keywords: cost-effectiveness analysis, decision rules, public health

1. Introduction

The use of formal methods of evaluation to inform adoption and reimbursement decisions for health technologies has become well established. However, there is a much wider set of policy decisions, including the implementation of public health interventions and programmes which could be informed by such analysis. For example the UK National Institute for Health and Clinical Excellence (NICE) has been given additional responsibilities for issuing guidance on public health interventions and national policies.[1 2] One of the features of this type of guidance is that it is likely to have an impact across public sector budgets and the wider economy. This poses the question whether the existing approach to the evaluation of technologies within the health care sector is sufficient to inform decisions across budget holders with multiple objectives and constraints. The general approach to economic evaluation in health care, and health technology assessment (HTA) in particular, can be characterised essentially as acceptance of a simple objective of maximising health gains (using some generally accepted pragmatic measure of health gain) subject to a single exogenous budget constraint. This approach can be seen as a species of 'extra-welfarism'[3 4] since it represents significant departures from the customary welfare foundations of cost-benefit analysis and related methods.

Increasingly sophisticated methods of structuring decision problems, synthesising evidence from a variety of sources, characterising uncertainty surrounding decisions and establishing the value of acquiring more information to inform these decisions have recently been developed.[5 6] These are beginning to be used to inform specific decision making – most notably as part of the technology appraisal programme at NICE.[7] These developments have taken place in a somewhat sheltered HTA world where both the single objective and the constraint are taken to be exogenous and regarded as legitimate on the basis that they are given by politically legitimate social decision makers. However, the simplification of constraints and objectives which may be reasonable when analysis is limited to health care technologies may not be appropriate from the wider view of the policies and interventions that are candidates for scrutiny in public health and wider social policy. The question arises of whether the current societal decision making approach - used, for example, by NICE - is sufficient for these wider policy questions, or whether generalisations of decision rules are required.

The current debate about which methods of evaluation should be used to inform these more complex decisions generally offers three broad alternatives: a somewhat ill-defined presentation of disaggregated costs and consequences (CCA); cost-benefit analysis (CBA) conducted from a welfarist perspective and cost-effectiveness analysis (CEA) which is commonly based on a social decision making approach. Some have recommended that CBA provides the most general approach in these circumstances but, where a complete CBA is not possible, CCA can be regarded as a practical alternative.[8] We start with a brief outline of what is required to inform decisions in HTA and identify those methods that best meet these requirements. We then ask whether public health interventions are fundamentally different from HTA and consider whether any such difference lies in the nature of decisions that must be made, the character of the evidence available and the objectives or the constraints faced. We try to locate the fundamental choices and social value judgments that must be made in choosing between alternative approaches.

2. Informing decisions in HTA

2.1 The decisions

Two decisions need to be made in health care. First, given existing evidence, which interventions or strategies should be adopted for defined patient groups, for particular indications and in specific settings? Second, is the existing evidence a sufficient basis for such a decision or is there value in acquiring further evidence? If the answer to the latter question is positive, what type of evidence, what type of studies for which patient groups and how much evidence is required? Where a technology appears to be cost-effective but uncertainty is high and more evidence would be an advantage, it may be better to delay implementation until the additional evidence has been acquired. Whether adoption should be delayed depends on the costs of reversing a decision to adopt now and the impact of immediate adoption on the probable generation of further evidence.

2.2 Analytical requirements

The choice of appropriate methods for economic evaluation should be based on whether they can provide the information required to make these decisions. The first requirement of any decision is some claim of legitimacy, which will ultimately rest on the social authority of the body with the remit to make these decisions and the legitimacy of the scientific and social value judgments which are required. Therefore, appropriate methods must make these scientific and social value judgements explicit so that they are open to debate, alternative formulation and falsification. Only then can their legitimacy be scrutinised.

If they were to be as fully informed as possible, these decisions would require an estimate of the joint distribution of costs and outcomes, for all the alternative interventions or strategies available (i.e., not just alternative technologies but all the different ways existing technologies could be used), for the full range of possible patient groups that can defined and over an appropriate time horizon (the time horizon over which cost and benefits differ). Therefore, a means of explicitly structuring the decision problem so that all alternatives are compared and that enables the explicit extrapolation of cost and benefits overtime, between settings and patient groups, is required.

A characterisation of the uncertainty surrounding the estimates of expected costs and benefits and the uncertainty surrounding the choice between the range of alternatives available is also required. No single study can provide all the information for the parameters relevant to costs and effects, nor is any single study going to be the only source of evidence.[9] Therefore, a means of combining evidence from a variety of sources and of representing uncertainty is required. Even when direct evidence about a parameter of interest is available it will often come from different types of study which will require explicit consideration of potential biases and exchangeability. More often than not, however, indirect evidence about parameters of interest is also available (such as evidence on functions of parameters, evidence on surrogate endpoints or indirect and mixed comparisons of the alternative technologies).[10 11] Moreover, there will also be circumstances where neither direct nor indirect evidence for a parameter is available and formal elicitation of scientific value judgements will be required.[12]

The framework of analysis also needs to reflect the iterative nature of decisions and the accumulation of evidence over time: as new evidence becomes available the estimates of costs and effects need to be updated and decisions may have to be revised.[13] Finally, the decision rules applied to the joint distribution of costs and other consequences must be consistent with the stated objectives and constraints of the health care system. With a single budget constraint and some view of its shadow price, and an objective of maximising health gains (where the measure of health gain has been agreed), the decision to adopt an intervention can be addressed. The second decision of whether the existing evidence is sufficient or whether further evidence is required can also be formally addressed using value of information analysis.[14] Similarly, the opportunity costs of delay while further information is acquired can also be established.[15 16] In principle all these decisions can be made in a way which is consistent with the objectives and constraints of the health care system.

These requirements suggest that a Bayesian decision analytic framework is appropriate, such as that developed within HTA. It was such a characterisation of the requirements for decision making in HTA which informed the selection of methods recommended in the Reference Case for the NICE appraisal of health technologies.[7] The question is whether this approach is sufficient when addressing the broader questions posed by public health interventions which move beyond the narrower concerns of HTA?

3. Public health interventions

What is really different about public health interventions compared to HTA? The decision still relates to the choice of technologies or interventions for the promotion of the health of the public. Decisions are still to be evidence-informed. Judgments will still need to be reached about the adequacy of the evidence to support decisions. Granted these similarities, the question of whether a new mode of analysis is needed seems redundant. It seems not to be needed. But that easy conclusion assumes that the objectives, the constraints and the general quality of the evidence are also not altered in ways

that may demand a different 'reference case'. Addressing these three issues is, therefore, what concerns us in the remainder of this article.

3.1 The evidence

It seems plain that the mix of types of evidence for the evaluation of public health interventions will differ in a number of important respects from that customarily available in medical care. For example, within HTA, partly due to licensing requirements, RCT evidence of the relative effect of interventions is often available, at least for pharmaceuticals. This experimental evidence avoids selection bias within the study and provides some of the best opportunities to explore the counterfactual. However, even in HTAs with a relative wealth of RCTs in the evidence base, there remain problems of interpretation, such as the external validity of protocol driven licensing trials, the shortness of follow-up periods, the use of surrogate endpoints and the frequent presence of incomplete and inappropriate comparators. It is the use of Bayesian methods of synthesis which has allowed these issues to be addressed [17 18] as well as incorporating other observational evidence which is almost always available but often ignored.[19]

The evaluation of public health interventions will inevitably rest more on non-experimental evidence concerning effectiveness such as natural experiments, cross-section studies or studies that include longitudinal and panel data with appropriate statistical analysis to control for observed and unobserved heterogeneity. In turn, these will require an integration of the econometric and statistical modelling approaches that have been applied to these problems (e.g., instrumental variables[20]) within a broader evidence synthesis. It will pose some new issues for synthesis methods which have, in HTA, focused primarily, but not entirely, on synthesis of evidence from RCTs. These issues will include how to synthesis the range of evidence of incentives on behaviour and consumption which have health effects. However, these are not issues that require new forms of analysis. What they do require is increased attention to the characteristics of empirical research that already pose interpretation problems in the appraisal of medical technologies.

In HTA evidence directly linking an intervention to health outcomes is sometimes, though rarely, available. The link between the intervention and long-term outcomes must often be made through a chain of evidence, linking the intervention to surrogate or intermediate outcomes (such as changes in behaviour/consumption, or biological markers) and then linking the surrogates to ultimate health outcomes. Linking surrogates to long term health outcomes and resource use may become a more frequent research issue in appraising public health interventions than in conventional HTA but the principles of how to address this issue are essentially the same.

The range of alternative interventions or policies to be considered in public health will be larger than with personal health care technologies, with fewer head-to-head comparisons. Again, this issue is central in conventional applications of HTA where many of the comparisons that are appropriate to make are not in the RCT evidence, and many comparisons remain ones between an experimental technology and placebo. Reflecting this dearth of appropriate head-to-head trial evidence, indirect and mixed comparisons is increasingly used in HTA.[11] So, once more, the challenge posed by public health to existing evaluative methods is neither new nor intractable. In fact, one of the most interesting aspects of extending this type of analysis to public health will be the opportunity afforded to compare health technologies to alternative public health interventions or even to non-health sector options.

3.2 Objectives

A naïve view

Current decision rules postulate a single objective that is to be maximised subject to a single exogenous budget constraint. In essence, HTA simply boils down to providing the solution to a well-specified constrained optimisation problem. Within HTA, the characterization of the exogenous objective function (for example, maximizing health outcome) may have been somewhat naïve. The naïvete arises for at least two similar but quite distinct reasons. First, while some analysts require that measures of outcomes should be preference-based and developed using the axioms of expected utility theory, it is not at all clear that measures of HRQL actually do measure utility as conceived by expected utility theory. Therefore, analysis based on the standard QALY can claim no firm

foundations in welfare theory.[21] This may already constitute an indictment of HTA as conventionally practised and would, *a fortiori*, be damned in the evaluation of public health interventions. However, this maybe of little concern for those who are sceptical as to whether utility theory can provide either behavioural predictions or a suitable basis for welfare propositions about the good for society.

Second, even when welfare theory is rejected as a suitable basis for decisions in health care there are a number of other concerns. While no one denies that health gain is not the sole outcome of health care, it is easy to defend health maximisation as a dominant criterion for the adoption of technologies into the bundle of services provided by a collectively funded health care system. In the case of public health, however, it is possible that other outcomes may need to be explicitly addressed.. Even when health gain is considered a suitable objective it may be too crudely measured (for example, in terms of quality-adjusted life-years (QALYs) using broad preference-based measures of health-related quality of life (HRQL) such as the EQ5D[22]) and so fail to capture subtle yet valuable aspects of health outcome. Finally a fair distribution of health outcomes is also an explicit expectation of public health interventions, but this aspect of evaluation has only been perfunctorily embodied in HTA as customarily practised.

A broader view

One alternative to the 'naïve' approach is to go to the other extreme, for example to postulate the maximisation of individual subjective utility, using the standard axioms of welfare economics to identify Pareto-preferred changes as indicators of improvements in social welfare, where an improvement in social welfare is judged to have taken place if those who gain can compensate those who lose from the intervention. Compensation can be based either on market prices that represent the social value of alternative activities or, when they do not, on shadow-prices assuming a first-best world. Where markets do not exist, revealed preferences in surrogate or hypothetical valuation can be use to generate shadow prices. On the face of it, such an approach has many advantages: all outcomes matter in so far as they are valued by individuals (who are willing to offer or accept compensation); it seems to avoid explicit interpersonal social value judgements and the inherent apparent paternalism of specifying some exogenous objective function; it provides clear guidance on appropriate methods (how costs and benefits should be measured and what perspective should be taken); and it provides a clear definition of what is meant by efficiency. This much broader view of the evaluation of social policy claims legitimacy from its firm foundation in individual preferences and seems ideally suited to public health interventions and the evaluation of broader social policies that affect health. It is a view clearly expressed by Mark Pauly:

"[Public] Heath care programmes should be judged in the same way as any other proposed change: ie the only question is do they represent a potential Pareto improvement not do they improve health outcomes as measured in either physical units or health state utility. It is possible that a programme may increase the health of some but reduce the health of others. If those that gain health outcome can compensate those that lose health (measured by individual willingness to pay) then the programme may be a potential Pareto improvement even if the health outcomes overall are lower." [23]

Pauly's customary clarity exposes the implicit social value judgements required to accept this approach, and these strong normative prescriptions come at a price: the implicit values may not necessarily be shared by legitimate societal decision makers; they are certainly not universally accepted; and they seem to contradict some quite basic values which are widely held. This was clearly expressed by the author of St. Luke's gospel:

"And he looked up and saw the rich putting their gifts into the treasury and He saw a poor widow putting in two small copper coins [mites]. And he said, "Truly I say to you, this poor widow put in more than all of them: for they all out of their surplus put into the offering but she out of her poverty put in all that she had to live on." Luke 21, v1-4, NAS.

For Mark this is clearly nonsense: the rich gave much more than the widow (they would require or could offer more compensation). Presumably, he would justify this by observing that the widow was poor because she had made free choices over her life cycle income and consumption, including her decisions to participate in the labour market and her investment in human capital consistent with her rates of time preference! There is no reason to intervene: to do so would be to violate individual preferences and might even constitute a dangerous attack on freedom itself.

Mark or Luke?

Choosing between Mark or Luke comes down to the position taken on two issues. The first is whether the current distribution of income (and the associated distribution of opportunities to invest in human capital and participate in the labour market) is optimal. If not, and if redistribution is possible, then the appropriate level of compensation can, in principle, be estimated for each alternative distribution, but we would need to agree which one was optimal and then adjust all compensation including shadow prices for marketed and non-marketed goods. Second, the approach views society as a collection of individuals participating in perfectly competitive markets, maximising their subjective utility following the conventional economic axioms of rationality. When this view of society is applied to the presumed nirvana of a first-best neoclassical world, market prices do indeed represent the social value of alternative activities (and, when they do not, they can be shadow-priced assuming a first-best world). This may be regarded as either a peculiarly optimistic or complacent view of the world which denies or abstracts from issues such as fellowship, solidarity, discrimination, class, imperialism and history.

In addition, there are a number of important problems in the application of the welfare theory: the conditions of rationality and consistency required for individuals maximising their utility have been shown often to be violated;[24] and there is a problem of aggregating compensating variations across individuals;[25] the paradox of choice reversal with non marginal changes;[26] and the general problem of 'second best'.[27] The last of these has generally received very little attention, despite the well known and profound result that first best solutions in a second best world may move us away from a Pareto optimum rather than towards one. Since no one would argue that the world is first best, then even if the values implicit in welfare economic theory were acceptable its successful application in a second best world seems impossible.

It seems clear that the strength of the normative prescriptions for social choice which flow from welfare theory rest on the strength of the social value judgements (and hence their ideological content) which are implicit in them. The notion that a walfarist view can avoid the strong social value judgements required by a societal decision making approach cannot be sustained. The difference is that the social value judgements required and the basis of claims for legitimacy are explicit under 'extra-welfarism'[3] principles but all too often implicit under welfarism. Moreover, it is plain that the social decision makers whom analysts serve do not see the objectives of health care in the same way as advocates of welfarism, whereas extra-welfarism is specifically designed to accommodate the values of social decision makers.

If not the invisible fist?

If the broader view of outcome based on individual preferences expressed through the market (actual, surrogate or contingent) is rejected, then what are the alternatives? There seem to be two: (a) to maintain the still somewhat naïve and simple objective for which there is some consensus (i.e. health technologies and public health interventions exist to improve health as much as possible); or (b) to attempt the specification of a more complex objective. Doing the latter would require specification of the arguments to be included, the weights attaching to each and some means by which political legitimacy could be claimed – that is, a rationale independent of the predilections of analysts themselves. Unfortunately, it is not clear how to resolve any of these issues, and no single legitimate societal decision maker may exist to specify how the resolution ought to be made. However, in the absence of a Leviathan, it maybe useful to look to those institutions which have been given the remit, and therefore some form of legitimacy, to make social decisions about health care and other sectors which may be relevant to public health interventions. How the valuation of outcomes across these different sectors might be achieved in a way which is consistent with the relevant budget constraints is considered below. However, before moving directly to a relaxation of the assumptions that the objective function has but one argument and that there is but one budget constraint, we consider the case of a single health sector with and without an exogenous constraint.

3.3 The constraints

A perfect institutional world

Suppose that the decision maker can set both the budget and allocate resources within it. Any project which improves social welfare as conventionally defined is to be implemented and the budget for health care will be determined by the resources needed to implement all projects with positive net benefit in CEA or positive net present value in CBA. The appropriate decision rules for CEA are quite clear, implement a new technology if:

$$\Delta c/\Delta H < V^H, \ or \ \Delta H.V^H - \Delta c^H > 0, \quad (1)$$

where health gain is the only benefit (validly measured by ΔH) and valued based on some social value of health gain (V^H). The additional cost of the technology (Δc) includes any net costs falling inside and outside the health care budget.

The decision rules in CBA are also quite clear, implement the new technology if:

$$\Delta c/\Delta B < 1, \ or \ \Delta B - \Delta c > 0, \quad (2)$$

where the benefits (ΔB) are valued based on individuals' maximum willingness to pay and Δc include all shadow priced costs, not just those falling on the health care sector.

Both CEA and CBA require a monetary valuation of health outcome. The distinction is the source of these valuations (individual willingness to pay versus some social value of health). Clearly, and as discussed previously, a CEA must ensure that all aspects of outcome that are valuable are included in ΔH. This may be difficult within HTA when the outcomes are all in terms of health gain but when applied to multiple sectors with diverse outcomes it is not clear what should be measured, how it should be measured and how these diverse outcomes should be valued. However, if the social value judgements required by a welfarist approach are acceptable, then CBA appears much more general (all outcomes are included if individuals are willing to pay for them) as long as the budgets for all other sectors are not fixed and the decision maker has the ability to allocate across as well as within all sectors.

A more realistic world

In practice, decision makers with the remit to allocate resources within sectors generally do not also have a remit to set their budget and allocate resource across sectors. A more realistic world is one of exogenous budget constraints. The decision rules for CEA in health care are again quite clear,[28] implement the new technology if:

$$\Delta c^H / \Delta H < \lambda, \ or \ \Delta H.\lambda - \Delta c^H > 0, \quad (3)$$

where λ represents the $\Delta c^H/\Delta H$ of the health care that will be displaced by the additional costs of the new technology (1/ λ represents the shadow price of the budget constraint). If the existing budget allocation is regarded as not just exogenous but also legitimate and in a relevant sense socially optimal then $\lambda=V^H$. Existing budgets can only be said to be suboptimal by making reference to some social value of health outcome, i.e., we can regard the health care system to be under-funded if $\lambda<V^H$ and if all costs are health care costs (Δc^H). However, allocation decisions within a single sector with a single objective and constraint do not require reference to an estimate of V^H because the same allocation decisions would be made irrespective of the value of V^H, provided that it is positive.

Appropriate decision rules using CBA are now less clear. Simply observing that the cost benefit ratio is less than 1 or that net present value is positive is no longer sufficient. We need to know whether the benefits offered by the new technology exceed the benefits which will be displaced elsewhere in the health care system. The technology should be implemented only if:

$$\Delta c^H / \Delta B < \gamma, \ or \ \Delta B - \Delta c^H / \gamma > 0, \quad (4)$$

where γ represents the cost benefit ratio of the health care which will be displaced and 1/ γ represents the shadow price of the budget constraint. Now we can regard the health care system to be under funded if γ<1 and if $\Delta c = \Delta c^H$, assuming that the individual willingness to pay which generates ΔB is regarded as the appropriate social valuation of health outcome.

The apparent generality of CBA is realised only in a perfect political and institutional world with no explicit budget constraints. Once an exogenous constraint is introduced similar problems are faced by both CEA and CBA. In both cases we now need an estimate of the shadow price of the budget constraint. We also need to know which costs fall on the constraint, since these are the only costs which will displace other activities. Similarly other benefits, in the form of cost savings which do not accrue to the health care budget, cannot offset costs which do. For these reasons, both HTA in general and the reference case of the NICE methodological guidance[7] restrict the perspective for costs to those falling on the health and personal social services sectors. However, there are potentially important elements of cost (e.g. to patients and carers) and benefits (e.g., productivity gains) which may fall outside these sectors. NICE policy is that these should be presented in a non reference case analysis.

Unfortunately, the NICE guidance is vague about how these other costs and benefits will be balanced against the health care and PSS costs and health benefits, simply stating that they maybe 'taken into account'. This problem is generally dealt with in an unsatisfactory way in HTA because there are no mechanisms to internalise such external effects. Similarly, the existence of an exogenous constraint also undermines the rationale for shadow pricing in CBA since the relevant costs are those that actually fall on the constraint, not the shadow price costs. For example, the shadow price of a new pharmaceutical will be substantially lower than the price actually faced by the health care system (due to monopoly rights granted by patent protection). Public health interventions with substantial costs and benefits falling on many different exogenous budgets pose these problems more dramatically for both the societal decision making and welfarist approaches.

A still more complex world

The allocation problem with exogenous constraints can be examined more completely when formulated as a mathematical programme.[29] Initially consider a health sector alone with an objective of maximising health outcome (H) or the monetary value of heath benefit (B^H) by choosing treatment j within programme k for population i subject to a single budget constraint for health, C_H:

$$\max_{\Psi}\left(\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=1}^{I_k} H_{ijk}\, x_{ijk}\right)$$

$$\Psi = \left(x_{ijk}, i = 1\ldots I_k,\ j = 1\ldots J_k, k = 1\ldots K\right)$$

or

$$\max_{\Psi}\left(\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=1}^{I_k} B^H_{ijk}\, x_{ijk}\right)$$

$$\Psi = \left(x_{ijk}, i = 1\ldots I_k,\ j = 1\ldots J_k, k = 1\ldots K\right)$$

sto

$$\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=1}^{I_k} c^H_{ijk}\, x_{ijk} \le C_H$$

$$0 \le x_{ijk} \le 1 \qquad i = 1\ldots I_k,\ j = 1\ldots J_k, k = 1\ldots K$$

$$\sum_{j=1}^{J_k} x_{ijk} = 1 \qquad i = 1\ldots I_k, k = 1\ldots K \qquad (5)$$

The solution to the dual of this simple linear programme provides 1/λ, or 1/γ, which, if used in the simple decision rules described, would determine an optimal selection of health technologies. However, there are three problems. First, the information requirements for even this very simple problem are enormous, requiring knowledge of *all* the costs (falling on the constraint) and *all* the health outcomes (or monetary benefits) of *every* competing alternative treatment for *each* programme and for *every* definable population group. For this reason estimates of λ, or γ have in practice to be made based on incomplete and imperfect information. Second, even when attention is confined to the health care system, recognition of multiple constraints leads to much more complex solutions. For example, the introduction of a simple budget constraint *in each time period* means that decision rules based on λ, or γ are no longer adequate.[30] Third, once the costs and benefits outside health are recognised, then the allocation problem becomes much more complex.

For example, consider education with the objective of maximising educational outcomes (E) or the monetary value of education benefit (B^E) by choosing policy j within programme k for population i subject to a single budget constraint for education C_E. The allocation problem within education could be considered in isolation from the impact on health outcomes and costs, just as the allocation problem in health can ignore any impact on educational outcomes and costs. However, in principle, the allocation problems within these two sectors can be solved simultaneously accounting for their respective budget constraints (solving for their respective shadow prices, $1/\lambda_H$ and $1/\lambda_E$) and the effects of health technologies on education outcomes and costs and those of educational programmes on health outcomes and costs. This requires specification of either a single objective or, where outcomes are measured on some utility scale or in natural units, specification of the relative weight of educational outcome compared to health outcome (δ). If all benefits are valued in monetary terms using CBA, they may be regarded as commensurate (each are valued relative to the common numeraire of consumption) and the objective will be to maximise total monetary benefit:

$$\max_{\Psi}\left(\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=1}^{I_k}(H_{ijk}+\delta.E_{ijk})x_{ijk}\right)$$

$$\Psi=\left(x_{ijk}, i=1\ldots I_k,\ j=1\ldots J_k, k=1\ldots K\right)$$

or

$$\max_{\Psi}\left(\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=1}^{I_k}(B^H_{ijk}+B^E_{ijk})x_{ijk}\right)$$

$$\Psi=\left(x_{ijk}, i=1\ldots I_k,\ j=1\ldots J_k, k=1\ldots K\right)$$

sto

$$\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=1}^{I_k}c^H_{ijk}\,x_{ijk}\le C_H$$

$$\sum_{k=1}^{K}\sum_{j=1}^{J_k}\sum_{i=1}^{I_k}c^E_{ijk}\,x_{ijk}\le C_E$$

$$0\le x_{ijk}\le 1 \qquad i=1\ldots I_k,\ j=1\ldots J_k, k=1\ldots K$$

$$\sum_{j=1}^{J_k}x_{ijk}=1 \qquad i=1\ldots I_k, k=1\ldots K \tag{6}$$

This type of formulation can evidently be extended to any number of sectors including the wider economy. Benefits for each sector may be measured in any combination of natural units, a utility scale or a monetary metric. However, if all sector outputs are not commensurate and valued relative to some common metric such as consumption then each must be valued relative to each other. When the wider economy is included they must also be valued relative to consumption in some legitimate and socially acceptable way. Aside from the difficulty of validly measuring outcomes and specifying the relative social value of each, the more fundamental problem is posed by the informational

requirements. Full information of *all* the costs and *all* the benefits falling on *every* sector for *each* competing alternative intervention across *all* programmes and *all* population groups for every sector would be required (this is similar to the scale of the problem of second best). While the specification of these informational demands offers no practical solution to the problem of evaluating interventions with multiple effects and multiple constraints, it demonstrates quite clearly that welfarist CBA offers no simple solution (the information requirements are equally demanding).

A pragmatic world

An alternative to attempting to solve the whole allocation problem across all sectors simultaneously in order to evaluate a single intervention is to focus on the impact of the intervention at the margin of each sector. If the existing allocation between the public sector and the wider economy and across the public sectors is regarded as not just exogenous but also legitimate and in a relevant sense socially optimal then the outcomes (or monetary benefits) within each sector can be valued based on informed estimates of the shadow price of each of the budget constraints faced. The impact of an intervention can then be expressed in terms of the net benefits falling on each sector. For example, a proposed intervention (j) which has an impact on outcomes and costs in both health and education will generate additional net benefit in health (ΔNB_j^H) and in education (ΔNB_j^E) compared to current practice:

$$\Delta NB_j^H = \Delta H_j.\lambda^H - \Delta c_j^H, \; or \; \Delta B_j^H - \Delta c_j^H \big/ \gamma_j^H \quad (7)$$

$$\Delta NB_j^E = \Delta H_j.\lambda^E - \Delta c_j^E, \; or \; \Delta B_j^E - \Delta c_j^E \big/ \gamma_j^E \quad (8)$$

The intervention should clearly be implemented if $\Delta NB_j^H > 0$ and $\Delta NB_j^E > 0$, since the gains in H and E are greater than the outcomes which will be displaced in each sector, similarly the intervention should be rejected if $\Delta NB_j^H < 0$ and $\Delta NB_j^E < 0$ (more outcomes or monetary benefit would be displaced than gained in both sectors). However, in many circumstances net benefit will be positive in one sector and negative in another. In such cases, implementation can be based a simple compensation test where the compensation required by a sector is based on net benefit. This ensures that the valuation of outcomes (or monetary benefits) is consistent with the existing budget allocation between sectors. This is illustrated in Table 1 for 6 notional interventions having different payoff patterns.

Table 1. A simple compensation test

j	Health	Education		Decision	Compensation
1	$\Delta NB_j^H > 0$	$\Delta NB_j^E > 0$	$(\Delta NB_j^H + \Delta NB_j^E) > 0$	Accept	None required
2	$\Delta NB_j^H > 0$	$\Delta NB_j^E < 0$	..	..	$(0- \Delta NB_j^E)$ from H to E
3	$\Delta NB_j^H < 0$	$\Delta NB_j^E > 0$	..	..	$(0- \Delta NB_j^H)$ from E to H
4	$\Delta NB_j^H < 0$	$\Delta NB_j^E < 0$	$(\Delta NB_j^H + \Delta NB_j^E) < 0$	Reject	None required
5	$\Delta NB_j^H > 0$	$\Delta NB_j^E < 0$	..	..	H can't compensate E
6	$\Delta NB_j^H < 0$	$\Delta NB_j^E > 0$	..	..	E can't compensate H

Thus, intervention 2 could represent an improvement in the quality of school meals. The gains in educational outcomes may not offset the costs falling on education ($\Delta NB_j^E < 0$) but the improvements in health outcomes and the possible reductions in health care costs means that $\Delta NB_j^H > 0$. The policy will be worth while implementing if the health sector can compensate education for the loss in net benefit ($0-\Delta NB_j^E$) and still regard the intervention to be cost-effective. Similarly, intervention 3 could represent the use of Ritalin for Attention Deficit Hyperactivity Disorder or parent training programmes which may not be cost-effective when the impact in health outcomes and costs alone are considered but the improvements in educational outcomes and possible reductions in special education costs means that the education sector could compensate the health sector for the loss in net benefit. The issue is not whether any particular perspective (e.g., societal or health system) is to be preferred. Rather that a multi-sectoral perspective is considered in which estimates of the impact on the outcomes deemed to be relevant across each sector and the actual costs (not the shadow prices) falling on each constraint are quantified.

It should be clear that the compensation required cannot be based only on the costs falling on different sector. Compensation should account also for the impact on outcomes across sectors. By expressing compensation in terms of net benefit based on λ_H and λ_E, the valuation of outcome and resource use in each sector is consistent with existing budget allocations across sectors. Indeed, λ_H/λ_E represents the value of health relative to educational outcomes implied by the current budget allocation and the productivity of current health and education interventions, i.e., 'as if' existing budget allocations are socially optimal.

Public and the private sector

The approach can be generalised to any number of sectors having explicit constraints as well as to the wider economy. It can also be generalised to consider policy decisions made by bodies which do not hold their own budget but make decisions which impact on other budgets and the wider economy. For example, legislation about workplace safety will impose costs on employers in the wider economy but will also improve health outcomes and may reduce health care costs. Whether such a policy should be regarded as worthwhile depends on whether the health sector could compensate the private sector for the additional costs but still regard the policy as cost-effective. The question is: do the health improvements valued at their opportunity cost (the additional health budget which would be required to generate the same improvement in health) net of health care costs, exceed the costs outside the health sector? If so, it is more efficient to generate health through the proposed policy than by expanding the budget for health care. Conversely, where health care interventions provide benefits to the wider economy (such as gains in productivity), they may be regarded as worthwhile if the private sector could compensate health care sector for any loss of net benefit.

If the compensation required for each intervention is actually paid then the impact of a decision on the net benefits in other sectors will be internalised. However, there typically exist no mechanisms for such transfers. Moreover, the payment of compensation for each intervention would entail genuine transaction costs. A more pragmatic approach might be to record the compensation required between sectors for implementation decisions made over a budgetary period. The net compensation required could be used to inform budget allocation for the next period. For example, if interventions 2 and 3 in Table 1 were implemented, then the net compensation required may be very small. This would be sustainable only if net compensation was either a small proportion of total budgets or if making changes in budgets based on net compensation became a credible political option. If that latter was the case, a public health intervention having, say, substantial costs ($\Delta NB_j^H < 0$) but also substantial productivity benefits to the wider economy, could be implemented only if there were confidence that the next period's budget would reflect compensation needed from the private sector to the relevant public sector budgets.

Are budget allocations optimal?

The approach outlined above takes the existing allocation of resources to and within the public sector as, not only exogenous, but also in some relevant sense socially optimal (the shadow prices are equal to social value of outcomes so that $V^H=\lambda_H$ and $V^E=\lambda_E$), i.e., the shadow price of the constraints is used to infer the value of sector-specific outcomes relative to each other and to consumption. This may not be unreasonable if we believe that the political and institutional processes that led to these allocations are broadly legitimate and represent social values which may not be captured in other ways (e.g., through the market whether real, surrogate or hypothetical). However, if budget allocations are regarded as exogenous but *not* socially optimal then this poses questions regarding the basis upon which such a claim can be made and what would be optimal allocations. Both require an alternative and appropriate social valuation of outcomes. Of course, only an unflinchingly welfarist approach provides a clear prescription: the appropriate social valuation is the individual willingness to pay (not some average social valuation which imposes external value judgments on individual preferences). Since these valuations have already been used to generate ΔB_j^H and ΔB_j^E the decision rule for intervention j remains unchanged and should be implemented if:

$$\Delta NB_j = (\Delta B_j^H - \Delta c_j^H / \gamma_j^H) + (\Delta B_j^E - \Delta c_j^E / \gamma_j^E) > 0 \qquad (9)$$

This net benefit of the intervention across sectors accounts for both the shadow price of the fixed constraints and the value of the outcomes in each sector relative to consumption. Observing that

$\gamma_H/\gamma_E=1$ would suggest that current allocations are consistent with individual preferences, whereas $\gamma_E<1$ and $\gamma_H/\gamma_E >1$ implies that there should be a reallocation from health to education (or, if γ_H is also less than 1, that both sectors are under funded).

If social valuation of outcomes based on individual willingness to pay through real, surrogate or hypothetical markets is not regarded as acceptable for the reasons outlined in Section 3.2, but the claim of sub optimal budget allocation is still maintained (i.e. the legitimacy of the political and institutional processes leading to the allocations are not a sufficient basis for social optimality), then some other 'legitimate' social valuation of sector-specific outcomes relative to consumption (V^H and V^E) is required and implied. The decision rule for intervention j which has impacts on health and education outcomes and costs must be revised to reflect both the social value of outcomes and the shadow price of the budget constraints.[31] The intervention should be implemented if:

$$\Delta NB_j = V^H(\Delta H_j - \Delta c_j^H / \lambda_j^H) + V^E(\Delta E_j - \Delta c_j^E / \lambda^E) > 0, \qquad (10)$$

Recall that in (3) when considering a single sector, with a single objective and constraint, the value of V^H did not effect the allocation decisions within that sector (it is simply a question of technical efficiency). However, when there are impacts on multiple sectors with different outcomes and when budgets are not regarded as socially optimal, the social value of outcomes matters for whether an intervention should be adopted. For example, for intervention 2 in Table 1 $\Delta NB_j = (\Delta H_j.\lambda_H - \Delta c_j^H)+(\Delta E_j.\lambda_E - \Delta c_j^E)>0$ but it is possible that $V^H(\Delta H_j.-\Delta c_j^H/\ \lambda_H)+V^E(\Delta E_j.-\Delta c_j^E/\ \lambda_E)<0$, if both sectors are under funded and if education is more under funded than health, i.e., $1<V^H/\ \lambda_H<V^E\ /\lambda_E$. In fact social valuations of sector outcomes only matter when $\lambda_H/\lambda_E \neq V^H/V^E$. It is possible that budgets are not optimal (both sectors are under funded) but both sectors are equally under funded. In these circumstances measures of net benefit in table 1 which are based only on the shadow price of the constraints will lead to the same allocation decisions as using social valuations in (10).

Clearly, claims about optimal allocation within and between the sectors can only be made with reference to some social values, e.g., observing that $\lambda_H/\lambda_E=V^H/V^E$ would suggest that current allocations are consistent with the socially determined relative value of health compared to educational outcomes. Whereas if $V^E>\lambda_E$ and $V^H/V^E<\lambda_H/\lambda_E$ implies that there should be a reallocation from health to education and if V^H is also greater than λ_H then both sectors are under funded.

Imposed or implied social values?

As discussed previously in Section 3.2 no single legitimate social decision maker exists to provide these social valuations and, if Mark is rejected in favour of Luke, it is not clear how they should be specified. However, any claims made about the optimality or otherwise of existing budget allocations appeals to such social values. The choice for analysts seems to be to either impose a view of social value (whether based on the market or some imaginary Leviathan) or evaluate interventions 'as if' existing allocation reflect social values. The former is much more ambitious (possibly to the point of being over reaching), the latter is more humble, but at least exposes the policy implications of and social values implicit in existing budget allocations made by those who claim some social legitimacy to make these decisions. This approach might have earned Keynes' approval:

"If economists could manage to get themselves thought of as humble, competent people, on a level with dentists, that would be splendid!" (p.373)[32];

4. Conclusions

There are pragmatic as well as principled reasons why a multi-sectoral societal decision making approach offers advantages. In the absence of a perfect political and institutional world in which first-best allocations are possible, exogenous constraints exist and decisions must be based on some assessment of their shadow price. There are, however, no simple ways of integrating the impacts on costs and benefits in other sectors within existing decision rules. The information requirements for formally solving the whole allocation problem across all sectors, accounting for multiple effects and constraints, are excessively demanding. Welfarist CBA offers no solution (the information requirements are the same) even if the implicit social value judgements are regarded as acceptable.

Unfortunately, if they are not acceptable, it is also the case that no single legitimate societal decision maker exists to specify the arguments and relative weights which would form some overall objective function.

However, if objectives and a measure of outcome for each sector can be based on institutions with a legitimate remit to make social decisions about allocations within each particular sector, then a pragmatic solution of the sort we have developed exists. Similarly, rather than impose some relative valuation of these outcomes, we can use the valuation implied by the current allocation of resources within the public sector and between the private and public sector, i.e. using the allocation decisions of those who have the remit and sufficient legitimacy to make them. This does not mean that the analyst must regard the status quo as in any sense 'correct' or 'optimal' but should conduct evaluation as if it is, given the explicit objectives, existing allocations and their implied values. By making these implications explicit, the question of whether existing objectives are appropriate and whether the implied valuation of outcomes across sectors matches social preferences is exposed to public scrutiny. In all cases the legitimacy of any prescription will rest on the legitimacy of the decision makers, the processes and the institutions. Explicitness and transparency about what current budget allocation decisions imply will serve as proximate tests of whether any claims for legitimacy are well-placed.

References

1. National Institute for Health and Clinical Excellence. *The public health guidance development process: An overview for stakeholders including public health practitioners, policy makers and the public*. London: National Institute for Health and Clinical Excellence, 2006.

2. National Institute for Health and Clinical Excellence (NICE). *Methods for Development of NICE Public Health Guidance*. London: NICE, 2006.

3. Culyer AJ. The normative economics of health care finance and provision. *Oxford Review of Economic Policy* 1989;5:34-58.

4. Hurley J. An overview of the normative economics of the health sector. In: Culyer AJ, Newhouse JP, editors. *Handbook of Health Economics*. Amsterdam: Elsevier, 2000.

5. Claxton K, Sculpher M, Drummond M. A rational framework for decision making by the National Institute for Clinical Excellence. *Lancet* 2002;360:711-715.

6. Sculpher M, Claxton K, Akehurst R. It's just evaluation for decision making: recent developments in, and challenges for, cost-effectiveness research. In: Smith PC, Ginnelly L, Sculpher M, editors. *Health Policy and Economics. Opportunities and Challenges*. Maidenhead: Open University Press, 2005.

7. National Institute for Clinical Excellence (NICE). *Guide to the Methods of Technology Appraisal*. London: NICE, 2004.

8. Kelly M, McDaid D, Ludbrook A, Powell J. *Economic Appraisal of Public Health Interventions - A Briefing Paper*. London: Health Development Agency, 2005.

9. Sculpher MJ, Claxton MJ, Drummond MJ, McCabe C. Whither trial-based economic evaluation for health care decision making? *Health Economics* 2006;15:677-687.

10. Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter evidence synthesis and consistency of evidence. *Statistics in Medicine* 2003;22:2995-3016.

11. Ades AE, Sculpher MJ, Sutton A, Abrams K, Cooper N, Welton N*, et al.* Bayesian methods for evidence synthesis in cost-effectiveness analysis. *Pharmacoeconomics* 2006;24:1-19.

12. O'Hagan A, Buck CE, Daneshkhah A, et al. *Uncertain Judgements: Eliciting Experts' Probabilities*. Chichester: John Wiley and Sons Ltd, 2006.

13. Fenwick E, Claxton K, Sculpher M, Briggs A. *Improving the efficiency and relevance of health technology assessment: The role of decision analytic modelling. Centre for Health Economics Discussion Paper 179*. York: Centre for Health Economics, University of York, 2000.

14. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Economics* 1996;5:513-524.

15. Eckermann S, Willan A. Expected value of information and decision making in HTA. *Health Economics* in press.

16. Griffin S, Claxton K, Palmer S, Sculpher MJ. Dangerous omissions: the consequences of ignoring decision uncertainty. *Health Economics* In submission.

17. Woolacoot N, Hawkins N, Mason A, Kainth A, et al. *Efalizumab and etanercept for the treatment of psoriasis. Assessment Report for NICE*. York: University of York, 2005.

18. Ward BG, Lloyd Jones M, A. P, et al. *Technology assessment report commissioned by the HTA Programme on behalf of the National Institute for Clinical Excellence. Statins for Prevention of Coronary Events.* London: NICE, 2005.

19. Sutton A, Abrams K, Jones D. Generalised synthesis of evidence and the threat of publication bias: the example of electronic fetal heart rate monitoring (EFM). *Journal of Clinical Epidemiology* 2002;55:1013-1024.

20. McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality. Analysis using instrumental variables. *Journal of the American Medical Association* 1994;272(11):859-866.

21. Pliskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. *Operations Research* 1980;28(1):206-224.

22. Kind P. The EuroQoL instrument: an index of health-related quality of life. In: Spilker B, editor. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia: Lippincott-Raven, 1996:191-201.

23. Pauly MV. Valuing health benefits in monetary terms. In: Sloan FA, editor. *Valuing Health Care. Costs, Benefits and Effectiveness of Pharmaceuticals and Other Medical Technologies*. Cambridge: Cambridge University Press, 1995.

24. Machina MJ. Choice under uncertainty: problems solved and unsolved. *Economic Perspectives* 1987;1:121-154.

25. Boadway RW. The welfare foundations of cost-benefits analysis. *Economic Journal* 1974;84:96-39.

26. Arrow K, Scitovsky T. *Readings in Welfare Eonomics*. London: Allen and Unwin, 1969.

27. Ng YK. *Welfare Economics: Introduction and Development of Basic Concepts*. London: MacMillan, 1983.

28. Weinstein MC, Zeckhauser R. Critical ratios and efficient allocation. *Journal of Public Economics* 1973;2:147-157.

29. Stinnett AA, Paltiel AD. Mathermatical programming for the efficent allocation of health care resources. *Journal of Health Economics* 1996;15:641-653.

30. Epstein D, Chalabi Z, Claxton K, Sculpher MJ. Efficiency, equity and budgetary policies: informing decisions using mathematical programming. *Medical Decision Making* in press.

31. Gravelle H, Brouwer W, Niessen L, Postma M, Rutten F. Discounting in economic evaluations: stepping forward to optimal decision rules. *Health Economics* in press.

32. Keynes JM. *Essays in Persuasion*. London: MacMillan, 1931.