ELSEVIER

Social Science & Medicine 62 (2006) 2091–2100

SOCIAL SCIENCE & MEDICINE

www.elsevier.com/locate/socscimed

Incremental cost-effectiveness ratios (ICERs): The silence of the lambda

Amiram Gafni, Stephen Birch*

McMaster University, Hamilton, Ont., Canada

Available online 1 December 2005

Abstract

Despite the central role of the threshold incremental cost-effectiveness ratio (ICER), or lambda (λ), in the methods and application of cost-effective analysis (CEA), little attention has been given to the determining the value of λ. In this paper we consider 'what explains the silence of the λ'? The concept of the threshold ICER is critically appraised. We show that there is 'silence of the λ' with respect to justification of the value of ICER thresholds, their use in decision-making and their relationship to the opportunity cost of marginal resources. Moreover, the 'sound of silence' extends to both 'automatic cut-off' and more sophisticated approaches to the use of λ in determining recommendations about health care programs. We argue that the threshold value provides no useful information for determining the efficiency of using available resources to support new health care programs. On the contrary, the threshold approach has lead to decisions that resulted in increased expenditures on health care programs and concerns about the sustainability of public funding for health care programs without any evidence of increases in total health gains. To improve efficiency in resource allocation, decision-makers need information about the opportunity costs of programs.

Keywords: Economic evaluation; Cost-effectiveness analysis; Decision rules; Incremental cost-effectiveness ratio; Health care costs

Introduction

Cost-effectiveness analysis (CEA) is presented in the research literature as a methodology to help decision-makers allocate scarce resources. The underlying premise of CEA is that for a given level of resources available, society or the decision-maker wishes to maximize the total aggregate health benefit conferred (e.g., Gold, Siegel, Russel, & Weinstein, 1996; National Institute for Clinical Excellence, 2004; Weinstein & Stason, 1977). The analytical tool of CEA is the incremental cost-effectiveness ratio (ICER) given by the difference in costs between two health care programs divided by the difference in outcomes between the programs with the comparison typically being between a new health care program and the existing approach to dealing with the same patient group.

The ICER has taken on increasing importance in the design, execution and use of cost-effectiveness research. In terms of the use of cost-effectiveness studies, the ICER provides a measure of average cost per additional QALY produced for use in deciding whether a new programme should be adopted. A decision to adopt or not adopt the program may be taken in terms of the program's

*Corresponding author.
E-mail addresses: gafni@mcmaster.ca (A. Gafni), birch@mcmaster.ca (S. Birch).

0277-9536/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2005.10.023

ICER value in relation to ICER values of competing programs, with health care resources being allocated in a descending order of ICER (i.e., first to programs with lower costs per QALY) until all available resources are exhausted (the 'league table' approach). Individual programs can also be judged in terms of the absolute value of the ICER. Programs with ICERs that lie below a 'threshold' ICER, also referred to as the lambda (λ) value, are deemed to be cost effective and should be adopted because the 'price' for producing health improvements implied by the ICER is acceptable (the 'threshold' approach).

The use of the threshold approach has formed the basis for recent developments in the methods for incorporating uncertainty into CEA. The net health benefit (NHB) approach (e.g., Stinnett & Mullahy, 1998) involves expressing incremental benefits and costs in the same units in order to calculate the difference term. But this requires knowledge of the specific value of λ. An alternative approach, involving cost-effectiveness acceptability curves (CEACs) (Briggs, 1999; van Hout, Al, Gordon, & Rutten, 1994) was developed in response to the value of λ being unknown. Under this approach information is provided on the probability that an intervention is cost effective for a given value of λ. In other words, it requires a decision-maker to judge what is an appropriate value of λ based on a range of possible λ values for which the intervention is 'cost effective' with a specific level of probability. Under either approach, a value for λ is required before a decision can be made.

Finally, there is an increasing tendency for economic evaluations to be performed prospectively as part of randomized controlled trials of new health care programs. This supports the use of conventional principles of statistical inference to quantify uncertainty in estimating ICER values. But this requires sample size calculations based on the requirements for the economic analysis, as well as those for the clinical trial. Briggs and Gray (1998) show how the sample size can be calculated for economic trials. However, this requires information on the value of λ. Willan and O'Brien (1999) consider the case where λ is a stochastic variable (as opposed to deterministic as under the Briggs and Gray approach). The required sample size is the one that produces a confidence region around the threshold value in the cost-effectiveness plane. However, this still requires information about the value of λ.

Despite the central role of the threshold ICER in the methods and application of CEA, little attention has been given to determining the value of λ. In this paper we consider 'what explains the silence of the λ?' The concept of the threshold ICER (λ) is critically appraised. We argue that the threshold value provides no useful information for determining the efficiency of introducing health care programs. On the contrary, the threshold approach leads to decisions that result in increased expenditures on health care programs without any evidence of increases in total health gains.

The theoretical foundation of the threshold ICER, λ

Weinstein and Zeckhauser (1973) consider the case of a government agency, working with a fixed budget, choosing between many projects, not all of which can be funded. They show that the 'critical ratio', λ, represents the opportunity cost of the resources at the margin. Under conditions of (i) perfect divisibility and (ii) constant returns to scale of all programs, which the authors assumed "… to avoid the problem of indivisibilities", they show that maximizing health benefits produced from available resources will occur under either of the following processes:

(a) all projects are ranked from the lowest to the highest ICER and selected in descending order until the resources are exhausted (the league table approach), or
(b) specification of the 'critical ratio', λ, directly and implementation of all projects with an ICER below or equal to λ (the threshold ICER approach).

There are several important implications of this theoretical approach to determining λ:

The health care budget: Under the Weinstein and Zeckhauser model, the critical ratio, λ, is a function of inter alia the size of the budget (Birch & Gafni, 1993). In other words, two communities with identical populations and faced with the same range of possible health care programs but different health care budgets will have different values of λ against which to judge the acceptability of programs (see Birch & Gafni, 2003). Similarly, every change in health care expenditure will generate a change in λ, ceteris paribus.

Uncertainty in the critical ratio: Because λ represents the opportunity cost of the marginal

health care resources it is equal to the ICER of the last program selected before the budget is exhausted. But the costs and effects of all programs including the last program selected for funding are subject to uncertainty. As a result λ is stochastic (see Sendi, Gafni, & Birch, 2002).

The dynamic nature of the critical ratio: The range of programs is dynamic, with new programs being developed over time. As new programs are funded and others replaced, the identification of the last program funded changes, implying that the distribution of λ also changes (Sendi et al., 2002).

Below we show that these properties of λ are not recognized in the applied cost-effectiveness literature and identify the implications of the use of a fixed λ, which does not represent the opportunity cost of marginal health care resources (i.e., the benefits forgone of the last unit of health care resources), for the efficiency in the use of health care resources.

From theory to practice: determining the value of λ

Because information on the incremental costs and effects of all current and potential programs is incomplete, the comprehensive league tables required to determine λ cannot be produced.[1] Restricted league tables based on those programs for which information is available could be used, but the ICER of the marginal program to be funded will not be a valid measure of λ. As a result, using the ICER of the last program funded under a restricted league table will not lead to the maximization of health improvements from available resources. Hence, regardless of whether one is willing to accept the theoretical assumptions of the model, the value of λ cannot be determined from the information available to the decision-maker (Devlin, 2002; Gafni & Birch, 2003a; Gold et al., 1996; Ubel, Hirth, Chernew, & Fendrick, 2003). No alternative approaches have been presented for determining λ, either in ways that are consistent with the Weinstein–Zeckhauser (1973) model, or based on any alternative theoretical models of health maximization from a constrained budget. However this has not prevented researchers claiming to identify the 'cost-effectiveness' of new programmes based on some 'preferred' or assumed value of λ.

For example, Laupacis, Feeny, Detsky, and Tugwell (1992) justified their choice of $20,000 (Canadian) per QALY "following a review of available economic evaluations and previously suggested guidelines". They argued that programs with ICERs less than this 'critical value' "…are almost universally accepted as being appropriate ways of using society's and the health care system's resources". No attempt is made by the authors to justify this figure in terms of it representing the marginal opportunity cost of health care expenditures in Canada in 1992. On the contrary, the ratio is derived from a 1982 US study, which suggested an arbitrary cut-off for programs at $20,000 (US) per QALY. Far from helping decision-makers determine whether funding a particular programme increases the health benefits produced from available resources, the use of similar arbitrary figures led to the allocation of *unconstrained* resources among programs without any evidence that overall health benefits were maximized (Gafni & Birch, 2003a,b; Laupacis, 2002). In particular, wherever the incremental cost of the new program is positive, increased resources are required to implement the program in place of the existing program no matter how large the incremental benefits are. Hence implementing new programs based on some acceptable λ, by definition, leads to an increase in health care expenditures.

Elsewhere US$50,000 to US$100,000 per QALY is often presented as a range for λ without any justification in terms of the opportunity cost of marginal health care resources, or the compatibility of its use with the maximization of health benefits from available resources (Ubel et al., 2003). The lower limit of this range is based on the ICER for renal dialysis treatment for patients with chronic renal failure although as Winkelmayer, Weinstein, Mittleman, Glynn, and Pliskin (2002) note "… it was initially expressed in Canadian rather than US dollars". Because in the US, Medicare is required by law to cover the cost of renal dialysis for all US citizens receiving the procedure, it has been argued that this represents a threshold that has been deemed to be an acceptable price to pay for health improvements in the US population. Hence, all interventions with ICER values less than or equal to this should be funded. However, Medicare does not fund other programs for all US citizens irrespective

[1] Other problems have been identified associated with the use of league tables by decision makers concerning, inter alia, the internal consistency of the cost effectiveness ratios within a table and the transferability of the data on which league tables are based between settings (see, for example, Donaldson & Gerard, 2005; Drummond, Torrance, & Mason, 1993).

of whether the ICER values are greater or less than this arbitrary figure. There may be many programmes that meet this critical ratio, but to fund them all would imply that the opportunity cost of health care resources was constant over whatever range of expenditures are required to support all these programs. In other words, it implies that at the extreme, there is an infinite stream of resources available at a constant marginal opportunity cost.

The use of similar ratios across settings involving different currencies further illustrates that the λ values bear no relation to the opportunity cost of marginal health care resources. For example, Nathoe et al. (2003) use the US$20,000/QALY as a threshold without reference to Laupacis et al. On the other hand, Oostenbrink et al. (2002) used 20,000 to 100,000 Euros but reference Laupacis et al. (1992) as the source.

Ubel et al. (2003) have raised concerns about the apparent constant value of the threshold over time. In particular they wonder why the value of the threshold has not increased in order to reflect the effect of inflation. They argue that the "acceptable threshold" of US$50,000/QALY to $100,000/QALY is too low and may have contributed to clinicians' discomfort with CEA. While acknowledging that "… there is no simple way to determine the appropriate price of a QALY", they argue that a much higher threshold value (i.e., US$265,000/QALY), would be more consistent with societal willingness to pay for medical interventions. However, no attempt is made to justify this, or any other value, in terms of λ as the opportunity cost of marginal health care resources. Indeed, the authors acknowledge the use of a threshold ICER to decide on which interventions are 'cost effective' would lead to continual increases in per capita health care costs, a somewhat strange outcome of an exercise aimed at providing the maximum health benefits from available resources as opposed to providing increased resources for producing benefits.

The US panel on CEA (Gold et al., 1996) acknowledge that no absolute standard exists for deciding whether a specific ICER value represents a cost-effective use of resources or not. Instead, they recommend describing programs as more or less cost effective based on the relative value of ICERs. However, they do not explain how this approach helps decision-makers allocate a particular pool of health care resources, whether that be the current budget or some other proposed or desired budget, in order to maximize health gains. As Doubilet, Weinstein, and McNeil (1986) noted, there is no theoretical justification that the strategy with the lowest cost-effectiveness ratio is the most desirable strategy from an economics perspective. For example Ubel et al. (2003) question why sildenafil has not been included as a treatment of erectile dysfunction on the formularies of various health care programs despite its very attractive ICER value of $11,000 per QALY. But introducing sildenafil would increase total health care costs, and not by only $11,000. Instead, providing the drug to all patients with erectile dysfunction would lead to a substantial increase in resource requirements that would require either major increases in health care resources or major cuts in other health care expenditures. Neither scenario represents an efficient use of resources a priori.

Hoch, Briggs, and Willan (2002) note the crucial role of λ in determining solutions to the constrained maximization problem facing decision-makers but suggest that this may be overcome "… if the decision maker can be assumed to know λ". But how do decision-makers determine λ? Have they developed or discovered a scientific approach that does not require information on the incremental costs and benefits of all programs? Do they have a solution to the problems of indivisibilities and non-constant returns to scale in programs? Or is this simply a convenient (albeit invalid) way for analysts to deal with the problem they are unable to solve for themselves?

Briggs, Goeree, Blackhouse, and O'Brien (2002), in an analysis of alternative approaches for the treatment of gastroesophageal reflux disease (GERD), acknowledge that the opportunity cost of a week free of GERD is not known, having already noted that if this were known "… it would be possible to choose between all the treatment strategies". Faced with this problem they continue with the analysis by adopting "… acceptable values".

Rawlins and Culyer (2004) explain that different thresholds are used under the National Institute for Clinical Excellence (NICE) guidelines in the UK. "Rather than apply an arbitrary threshold, … decisions (are) based on a case by case basis… As the incremental cost effectiveness ratio increases, the likelihood of rejection on grounds of cost effectiveness rises". Interventions with ratios below a lower threshold would generally be regarded as cost effective and recommended for adoption without further consideration. Those with ratios above a

higher threshold would only be recommended for adoption under particular circumstances (e.g., important equity considerations). However no empirical basis is provided for the assigned ranges of values of £5,000–£15,000 and £25,000–£35,000 for the lower and upper thresholds, respectively. Neither is the use of these (or other) thresholds justified in terms of making best use of NHS resources. In particular, in making judgements about the cost effectiveness of interventions the NICE guidelines are applied without taking into account affordability (Rawlins & Culyer, 2004). However, as Williams (2004) notes, such considerations are the only justification there is for having thresholds. He acknowledges the fact that there is no practical way to determine the threshold and thus suggests "a bit of common sense". He argues that in the UK there are £18,000 worth of real resources per citizen to provide for all needs (e.g., food, shelter, transport). He suggests adopting this figure as the threshold value because "it is clear that we could do that at the margin for a few people without imposing great hardships on the bulk of the population but we could not do it for many". He acknowledges that this threshold value is "just my opinion" and does not attempt to show how use of this threshold would be compatible with maximizing health gains from available resources, or any other objective for the NHS.

Cost-effectiveness thresholds and value for money in the allocation of health care resources: myth or reality?

In this section we discuss how the threshold value, λ, is being used and identify the consequences of using λ values of unknown or unsubstantiated origin? The main approach used to determine whether a program should be implemented employs λ as an 'automatic cut-off' decision rule based on the theoretical model described earlier (Birch & Gafni, 1992; Weinstein & Zeckhauser, 1973). The advantage of relating the threshold to the theoretical model is that if (i) the underlying assumptions of the model hold in real life and (ii) λ is determined in a way consistent with the model, the use of the this decision rule will guarantee that programs will be chosen in a way that maximizes health gains from available resources.

Methodological guidelines for the economic evaluation of health care programs based on the 'automatic cut-off' ICER threshold were developed and proposed for the adoption of new technologies in Canada over 10 years ago (Laupacis et al., 1992). Although at the time we explained that these guidelines represented a prescription for uncontrolled growth in drug expenditures (Gafni & Birch, 1993), the guidelines formed the basis for recommending which drugs should be covered by public funding for senior citizens and those on social assistance. The chair of the committee responsible for making recommendations, who was also principal author of the guidelines, notes that because "resources for health care are limited, it seems sensible…that cost-effectiveness is the main criterion used to determine which drugs are reimbursed from the public purse" (Laupacis, 2002). As explained above, the use of the ICER threshold approach requires increased expenditures on the program for every new drug introduced wherever the ICER value is positive (i.e., drug has greater effects but costs more than the current treatment). In the case of this drugs program, total costs increased by 10% per annum over the period 1997–2000 and by 15% in 2001. No information is available on what services were reduced and with what consequences for health in order to support these increased expenditures on the drug program. This uncontrolled growth in program expenditures led to questions being asked about the affordability and sustainability of the program by both the Premier of Ontario and his Minister of Health (Laupacis, 2002).

In Australia an ICER threshold approach was introduced for decision-making for the pharmaceutical benefits programme in the early 1990s as a means of controlling cost escalation and promoting efficient use of programme resources. Costs of the programme were observed to increase by over 14% per annum over the first 10 years of using the approach (Zinn, 2002). Again, the effect of these increased expenditures on drugs for the overall increases in health gains among Australian residents remains unknown.

In the UK, recommendations made by the NICE about the appropriateness of technologies for use under the National Health Service (NHS) are based on the calculated ICERs of the technologies. Once accepted by the Secretary of State for Health the recommendations become mandatory on health authorities. According to Taylor (2002) the first 10 drugs recommended for adoption by NICE required additional funding for the NHS of 250 million pounds. NICE was unable to provide an answer to

the Parliamentary Select Committee question about how much these recommendations added to actual NHS expenditures because of 'methodological difficulties'. However, other evidence reported that NHS expenditure on these drugs was around 60% of this figure (House of Commons' Select Committee, 2002). The discrepancy is apparently explained by the failure of some authorities to implement the NICE recommendations. For example 15–20% of Health Authorities could not confirm that they have introduced the recommended technologies 1 year after publication of the recommendation (Taylor, 2002) while fewer than half of all health authorities have a policy for monitoring local compliance with NICE recommendations. More recent estimates suggest the additional resources absorbed by NICE recommended technologies (and hence the additional revenues of those corporations involved in the production of those technologies) exceeded £575 m in the first 2.5 years of NICE (Mayor, 2002).

The presence of a NICE threshold was consistently denied by those involved with the Institute, until publication of the most recent guidelines (Devlin, Parkin, & Gold, 2003; House of Commons' Select Committee, 2001). Culyer (2002) argued that not having an explicit threshold may simply represent "a perfectly consistent application of a rather more sophisticated decision rule than an automatic cut-off value, or it may result from the application of a multiple-decision guidelines". In other words, ICER values are only one input into the decision-making process. Other inputs, equity considerations being prominent examples, may also be important to the decision-making process. For example, NICE is concerned with both (a) the maximization the health gain from the use of NHS resources and (b) removing unfairness in the availability of technologies (National Institute for Clinical Excellence, 2001a). Under the NICE guidelines, equity issues are to be addressed in three ways: the distribution of potential benefits from the intervention by clinical and social status; the distribution of personal costs on patients and their families; and the distribution of NHS costs by social groups. However, no information is provided on how these considerations are to be incorporated into the 'sophisticated decision rule' to be used to determine which technology should be adopted (National Institute for Clinical Excellence, 2001a).

The NICE appraisal of Riluzole, a drug used in the treatment of Motor Neuron disease, provides an interesting example. The study on which the NICE appraisal was based calculated the ICER for the drug to be between £34,000 and £43,000 per QALY (National Institute for Clinical Excellence, 2001b). Although NICE denied the existence of a threshold, a WHO review concluded that it "… was clear that NICE has a threshold and that it is £30,000 per quality adjusted life year (QALY)" (Devlin et al., 2003). No explicit justification was provided in the NICE appraisal to support the recommendation to introduce the drug, even though other technologies with similar ICERs were not recommended for implementation. However the appraisal did emphasize the severity of the condition and strong patient preferences for years of life free of tracheostomy as factors important in reaching its recommendation (National Institute for Clinical Excellence, 2001b). It seems that these factors were presented as considerations that justify weighting the denominator in the ICER by an amount sufficient to reduce the ICER value of the drug to the implicit threshold of £30,000 per QALY. However, these factors were already included in the calculation of QALYs. To use these factors to support the adoption of this drug implies double counting severity and lifespan considerations in arriving at the recommendation. Moreover, to (implicitly) weight the ICER based on *patients' preferences*, in addition to societal preferences that lay behind the QALY scores appears to be somewhat inconsistent when patient preferences are not considered as important information in other NICE appraisals.

As already mentioned the precise nature of the sophisticated decision rule is not explained. However, in the recently updated NICE guidelines (National Institute for Clinical Excellence, 2004) two separate thresholds are presented. Programs with ICERs below £20,000/QALY are recommended for adoption based primarily on the ICER value, i.e., without reference to any other factors. Programs with ICERs between £20,000/QALY and £30,000/QALY are more likely to require additional justification, such as the degree of uncertainty in estimated calculations, the innovative nature of the technology or features of the condition or population receiving the intervention, in order to be recommended by NICE. Programs with ICERs above £30,000/QALY will only be recommended for adoption in the NHS if the case for supporting the technology on these additional considerations is "increasingly strong" (National Institute for Clinical Excellence, 2004). No attempt is made to justify the different threshold values or to explain how the

application of the thresholds in conjunction with other components of the sophisticated decision rule, achieves the NICE goals.

In order to understand how guidelines might assist decision-makers in making difficult choices about the adoption of technologies, it is important that the guidelines be justified and that the justification be transparent (Birch & Gafni, 2002). For example, it is wrong to assume that the Weinstein–Zeckhauser model giving rise to the 'automatic cut-off' rule does not incorporate equity considerations. Efficiency involves maximizing an objective function that aggregates health benefits across individual patients, subject to a constraint. The process of aggregation requires the use of some rule about the relative weights to be attached to the distribution of benefits across patients. One may not like the equity position incorporated in the Weinstein–Zeckhauser model, where all health gains are weighted equally, irrespective of who receives them. In this case one can incorporate the preferred equity position into the objective function (e.g., to 'favour' outcomes accruing to one particular individual or social group) or incorporate it as an additional constraint (e.g., to ensure equal availability of services, irrespective of outcomes). By incorporating the 'additional' considerations into the underlying model of constrained maximization we ensure an explicit and systematic consideration of the opportunity costs of pursuing these considerations. As Williams and Cookson (2000) argue

> [I]f the nature and implication of … equity principles are to be clarified in a policy-relevant way, it is necessary to quantify what might otherwise merely remain vaguely appealing but ambiguous slogans. Only with some quantification will it be possible to convert them into criteria that can be applied in a consistent manner, and with a reasonable chance of checking on performance (i.e., holding people accountable).

In summary, there is no evidence to suggest that the application of the ICER threshold approach, either in the form of an 'automatic cut-off'' or as a component of a 'sophisticated decision rule', has helped decision-makers achieve the stated objectives, whether those objectives are simply the maximization of health gains produced in the population *per se* or involve some additional considerations. Instead there is a remarkable lack of attention devoted to the justification of the decision-making criteria adopted and the outcomes of those decisions. Although individual programs may lead to unambiguous improvements in health among patient groups for whom these programs are provided, the opportunity costs associated with these programs remain unknown and conceivably could lead to a reduction in the aggregate health effects from available resources. The only unambiguous conclusion that can be reached is that the use of ICER thresholds does increase total costs and hence the aggregate revenues of the manufacturers of new technologies.

Discussion

In this paper we have focused attention on the use of ICER 'threshold' approaches to decision-making about alternative uses of health care resources given the stated objectives of decision-makers about the maximization of health gains from available resources. It was not the purpose of this paper to challenge these objectives, either in terms of the relevance of health gain as the appropriate maximand, or the validity of instruments such as QALYs as measures of health gains although there are substantial literatures on both of these issues (e.g., Birch & Donaldson, 2003; Birch, Melnikow, & Kupperman, 2003; Gafni & Birch, 1995, 1997).

Despite the central role of λ in identifying the cost-effectiveness of health care programs, the determination of λ in practical applications of economic evaluation remains elusive. The theoretical methods cannot be applied because of the underlying assumptions of the model and incomplete data on the incremental costs and effects of programs. The use of 'acceptable' values for λ represents a departure from the concept of the opportunity cost of marginal resources. The values of λ are not related to the size of the available pool of resources. Moreover the λ values adopted are deterministic and constant and hence they fail to represent the uncertain nature of the last program funded as well as the changing identity of that program as the portfolio of programs changes. Finally, the use of 'acceptable' values for λ implies the availability of an indeterminate stream of additional resources at a constant marginal opportunity cost (Birch & Gafni, 1993).

The current literature, faced with the problem of incomplete information has looked towards specifying 'acceptable' thresholds. Authors rarely justify what makes the chosen values 'acceptable' and do

not consider the implications of using these values for the efficiency of use of available resources. This casts doubt on the usefulness of the methods developed for incorporating uncertainty into CEA and for sample size calculations for economic evaluations performed prospectively as part of randomized controlled trials since these require knowledge of the value of λ. Furthermore, it is not clear how these methods can be used where λ is just one component of a sophisticated decision rule.

Silence with respect to the value of λ and the practical constraints on its derivation are acknowledged (e.g., Devlin, 2002; Eichler, Kong, Gerth, Mavros, & Jonsson, 2004; Gafni & Birch, 2003a; Gold et al., 1996; Ubel et al., 2003). This reminds us of the decision science concept of the 'Phantom Alternative' represented by an illusionary choice. It looks real but for some reason is unavailable at the time a decision is made. On the one hand, phantom alternatives can provide useful information on the boundaries of a decision problem and thus help generate new options through a restructuring of the problem. But, as illustrated in this paper and recognized in the literature, phantoms can also produce biases, deception and sub-optimal decisions (Farquhar & Pratkanis, 1993).

So what is the role of economics in assisting decision-makers? Laupacis (2002) argues that the determination of 'cost-effectiveness' as a criteria for recommending adoption of a new drug onto the formulary "… is often determined by effectiveness not costs". Similarly, the Policy Advisory Group responsible for reviewing practice guidelines for the treatment of cancer in Ontario acknowledges using a threshold approach to decisions about introducing new drugs for the treatment of cancer. However, they concede that in terms of criteria for making recommendations "it is unlikely that a single metric, for example, cost per quality adjusted life years, would be satisfactory" (Pater et al., 2001). It is interesting that in both cases, the recommendations of the respective decision-making bodies have not been constrained by a given resource pool. In other words, the concept of the λ, the opportunity cost of marginal resources, was irrelevant to the decision-makers. Irrespective of whether the problem faced by decision-makers is simple (maximizing health gains from available resources) or complex (subject to considerations of equity, accessibility, etc.), if it is not to be considered in the context of a resource constraint there is little use for economics in the way the problem is considered. Resource consequences of the decisions must simply be accepted irrespective of the opportunity costs of these decisions. But this means that there is no reason to believe the decisions taken lead to available resources producing maximum health gains. Eichler et al. (2004) see explicit ICER thresholds emerging in the future in many countries as a means of improving the transparency and consistency of decision-making as if these where the objectives of decision-makers. But as we indicated above, the opportunity cost of marginal health care resources is a dynamic concept and its value will change as new programmes are funded and/or resource constraints change. Hence consistency in the values of λ used or implied from decisions is unlikely to be compatible with efficiency in resource allocations.

An alternative approach to the theoretical and practical problems of determining λ is to estimate λ based on measuring society's willingness to pay (WTP) for additional health outcomes. It has been argued that this would provide a clearer link between practical CEA and the theoretically superior cost–benefit analysis derived from welfare economic theory (Bleichrodt & Quiggin, 1999; Johannesson, 1995; O'Brien, Gertsen, Willan, & Faulkner, 2002). However, there is no reason to believe that WTP per QALY is constant over a wide range of QALYs. As with other commodities, we might expect that the marginal utility of QALYs diminishes with the size of QALY production. In addition, the opportunity cost of a QALY will increase with the size of QALY production as the available resources from which to fund additional QALYs diminishes. It is therefore difficult to see how information on WTP per QALY in a population sample could be used to determine how to allocate available health care resources in order to maximize health gains in the population. It seems far more plausible that such an approach might be used to 'justify' the increased expenditures associated with the failure of the existing economic evaluation guidelines to address this question.

Economics can be used to determine whether adoption of a new program represents an efficient use of health care resources and applies is generally to scenarios of fixed, shrinking or increasing budgets. To deal with situations of less than complete information on all possible programs, a modified approach has been presented to determine whether the adoption of a new program represents an unambiguous improvement in the efficiency of resource use (i.e., more health benefits from

available resources)(Birch & Gafni, 1992; Sendi et al., 2002). Under this approach direct comparisons are made between the incremental benefits associated with the new program and the incremental benefits associated with those programs that must be cancelled or reduced in order to generate the additional resources required by the new program. Because this involves the direct consideration of opportunity costs, measured in terms of health benefits forgone, it takes the form of a (non-monetary) cost–benefit analysis.

Program Budgeting and Marginal Analysis (PBMA) has been presented and adopted in some jurisdictions as an approach aimed at maximizing health gains from available resources (see Mitton & Donaldson, 2004) and incorporates many of the features of this modified approach. In particular, it aims to consider the opportunity cost of increasing resource allocations to particular programmes directly, in terms of which other programs will be reduced. However, the comparisons of benefits produced from expanding one program and benefits forgone in reducing another are made on the basis of the "greater benefit per pound spent" (Mitton & Donaldson, 2004), i.e., the ICER for each program. Hence, the approach faces many of the same problems of traditional ICER threshold approach (e.g., assumptions of constant returns to scale and perfect divisibility I al programmes).

We have shown that there is 'silence of the λ' with respect to justification of the value of ICER thresholds, their use in decision-making and their relationship to the opportunity cost of marginal resources. Moreover, the 'sounds of silence' extend to both 'automatic cut-off' and more sophisticated approaches to the use of λ in determining recommendations about health care programs. However, Cookson, McDaid, and Maynard (2001) note that

> To improve efficiency, decision makers need information on what economists call opportunity costs—the benefits forgone when scarce resources are used one way rather than another… In absence of any information about opportunity cost, however, they cannot attempt to achieve the efficient use of resources.

Failure to be efficient about the way we pursue stated objectives, whether they be concerned with maximizing the aggregate health benefits generated, equity in the distribution of health outcomes, equality in the availability of services or a combination of these aspirations, involves an unambiguous waste of scarce resources. We know of no branch of economics where this is an acceptable outcome.

References

Birch, S., & Donaldson, C. (2003). Valuing the benefits and costs of health care programmes: Where's the 'extra' in extra-welfarism? *Social Science & Medicine*, *56*, 1121–1133.

Birch, S., & Gafni, A. (1992). Cost effectiveness/utility analysis: Do current decision rules lead us to where we want to be? *Journal of Health Economics*, *11*, 279–296.

Birch, S., & Gafni, A. (1993). Changing the problem to fit the solution: Johannesson and Weinstein's (mis)application of economics to real world problems. *Journal of Health Economics*, *12*, 469–476.

Birch, S., & Gafni, A. (2002). On being NICE in the UK: Guidelines for technology appraisal for the NHS in England and Wales. *Health Economics*, *11*, 185–191.

Birch, S., & Gafni, A. (2003). Economics and the evaluation of health care programmes: Generalizability of methods and implications for generalizability of results. *Health Policy*, *64*, 207–219.

Birch, S., Melnikow, J., & Kupperman, M. (2003). Conservative versus aggressive follow-up of mildly abnormal Pap smears: Testing for process utility. *Health Economics*, *12*, 879–884.

Bleichrodt, H., & Quiggin, J. (1999). Life-cycle preferences over consumption and health: When is cost effectiveness analysis equivalent to cost benefit analysis? *Journal of Health Economics*, *18*, 681–708.

Briggs, A. (1999). A Bayesian approach to stochastic cost-effectiveness analysis. *Health Economics*, *8*, 257–261.

Briggs, A. H., & Gray, A. M. (1998). Power and sample size calculations for stochastic cost-effectiveness analysis. *Medical Decision Making*, *18*, S81–S92.

Briggs, A. H., Goeree, R., Blackhouse, G., & O'Brien, B. J. (2002). Probabilistic analysis of cost effectiveness models: Choosing between treatment strategies for gastroesophageal reflux disease. *Medical Decision Making*, *22*, 290–308.

Cookson, R., McDaid, D., & Maynard, A. (2001). Worng SIGN, NICE mess: Is national guidance distorting allocation of resources? *British Medical Journal*, *323*, 743–745.

Culyer, T. (2002). Introduction: Ought NICE to have a cost-effectiveness threshold? In A. Towse, C. Pritchard, & N. Devlin (Eds.), *Cost effectiveness thresholds: Economics and ethical issues* (pp. 9–14). London: Kings Fund and Office of Health Economics.

Devlin, N. (2002). An introduction to the use of cost-effectiveness thresholds in decision- making: What are the issues? In A. Towse, C. Pritchard, & N. Devlin (Eds.), *Cost effectiveness thresholds: Economics and ethical issues* (pp. 16–24). London: Kings Fund and Office of Health Economics.

Devlin, N., Parkin, D., & Gold, M. (2003). WHO evaluates NICE: The report card is good but incomplete. *British Medical Journal*, *327*, 1061–1062.

Donaldson, C., & Gerard, K. (2005). *Economics of health care financing. The visible hand* (2nd ed.). Basingstoke: Palgrave Macmillan.

Doubilet, P., Weinstein, M. C., & McNeil, B. J. (1986). Use and misuse of the term "cost effective" in medicine. *New England Journal of Medicine*, *314*, 253–256.

Drummond, M., Torrance, G., & Mason, J. (1993). Cost-effectiveness league tables: More harm than good? *Social Science & Medicine*, *37*, 33–40.

Eichler, H., Kong, S., Gerth, W., Mavros, P., & Jonsson, B. (2004). Use of cost-effectiveness analysis in health care resource allocation decision-making: How are cost-effectiveness thresholds expected to emerge? *Value in Health*, *7*, 518–528.

Farquhar, P., & Pratkanis, A. (1993). Decision structuring with phantom alternatives. *Management Science*, *39*, 1214–1226.

Gafni, A., & Birch, S. (1993). Guidelines for the adoption of new technologies: A prescription for uncontrolled growth in expenditures and to avoid the problem. *Canadian Medical Association Journal*, *148*, 913–917.

Gafni, A., & Birch, S. (1995). Preferences for outcomes in economic evaluation: An economic approach to addressing economic problems. *Social Science & Medicine*, *40*, 767–776.

Gafni, A., & Birch, S. (1997). QALYs and HYEs: Spotting the differences. *Journal of Health Economics*, *16*, 601–608.

Gafni, A., & Birch, S. (2003a). NICE methodological guidelines and decision making in the National Health Service in England and Wales. *Pharmacoeconomics*, *21*, 149–157.

Gafni, A., & Birch, S. (2003b). Inclusions of drugs in provincial drug benefit programs: Should 'reasonable decisions' lead to uncontrolled growth in expenditures? *Canadian Medical Association Journal*, *168*, 849–851.

Gold, M. R., Siegel, J. E., Russel, L. B., & Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine*. New York: Oxford University Press.

Hoch, J. S., Briggs, A. H., & Willan, A. R. (2002). Something old, something new, something borrowed, something blue: A framework for the marriage of health econometrics and cost effectiveness analysis. *Health Economics*, *11*, 415–430.

House of Commons' Select Committee. (2001). *Health: National Institute for Clinical Excellence*. London: Her Majesty's Stationary Office.

House of Commons' Select Committee. (2002). *Health: National Institute for Clinical Excellence*. London: Her Majesty's Stationery Office.

Johannesson, M. (1995). The relationship between cost effectiveness analysis and cost benefit analysis. *Social Science & Medicine*, *41*, 483–489.

Laupacis, A. (2002). Inclusion of drugs in provincial drug benefit programs: Who is making these decisions, and are they the right ones? *Canadian Medical Association Journal*, *166*, 44–47.

Laupacis, A., Feeny, D., Detsky, A., & Tugwell, P. X. (1992). How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Canadian Medical Association Journal*, *146*, 473–481.

Mayor, S. (2002). News extra: NICE estimates that its recommendations have cost the NHS £575 m. *British Medical Journal*, *325*, 924.

Mitton, C., & Donaldson, C. (2004). *Priority setting toolkit: A guide to the use of economics in healthcare decision-making*. London: BMJ Books.

Nathoe, H. M., van Dijk, D., Jansen, E. W. L., Suyker, W. J. L., Diephuis, J. C., van Boven, W. J., et al. (2003). A comparison of on-pump and off-pump coronary bypass surgery in low-risk patients. *New England Journal of Medicine*, *348*, 394–402.

National Institute for Clinical Excellence. (2001a). *Technical guidance for manufacturers and sponsors on making a submission to a technology appraisal*. London: National Institute for Clinical Excellence.

National Institute for Clinical Excellence. (2001b). *Riluzole for Motor Neurone disease—full guidance*. London: National Institute for Clinical Excellence.

National Institute for Clinical Excellence (2004). *Guide to the Methods of Technology Appraisal*. Available from URL: http://www.nice.org.uk/pdf/brdnov03item3b/pdf (accessed 13 January 2004).

O'Brien, B. J., Gertsen, K., Willan, A. R., & Faulkner, L. A. (2002). Is there a kink in consumers' threshold value for cost effectiveness in health care? *Health Economics*, *11*, 175–180.

Oostenbrink, R., Oostenbrink, J. B., Moons, K. G. M., Derksen-Lubsen, G., Essink-Bot, M. L., Grobbee, D. E., et al. (2002). Cost utility analysis of patient care in children with meningeal signs. *International Journal of Technology Assessment in Health Care*, *18*, 485–496.

Pater, J. L., Browman, G. P., Brouwers, M. C., Nefsky, M. F., Evans, W. K., & Cowan, D. H. (2001). Funding new cancer drugs in Ontario: Closing the loop in practice guidelines development cycle. *Journal of Clinical Oncology*, *19*, 3392–3396.

Rawlins, M., & Culyer, A. (2004). National Institute for Clinical Excellence and its value judgements. *British Medical Journal*, *329*, 224–227.

Sendi, P., Gafni, A., & Birch, S. (2002). Opportunity costs and uncertainty in the economic evaluation of health care interventions. *Health Economics*, *11*, 23–31.

Stinnett, A. A., & Mullahy, J. (1998). Net health benefit: A new framework for the analysis of uncertainty in cost effectiveness analysis. *Medical Decision Making*, *18*, S68–S80.

Taylor, R. (2002). Generating national guidelines—a NICE model? *Paper presented at the fifth international conference on strategic issues in health care management policy, finance and performance in health care*, St. Andrews, Scotland, UK, April 11–13.

Ubel, P. A., Hirth, R. A., Chernew, M. E., & Fendrick, A. M. (2003). What is the price of life and why doesn't it increase at the rate of inflation? *Archive of Internal Medicine*, *163*, 1637–1641.

van Hout, B. B. A., Al, M. J., Gordon, G. S., & Rutten, F. F. (1994). Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics*, *3*, 309–319.

Weinstein, M., & Zeckhauser, R. (1973). Critical ratios and efficient allocation. *Journal of Public Economics*, *2*, 147–157.

Weinstein, M. C., & Stason, W. B. (1977). Foundation of cost effectiveness analysis for health and medical practices. *New England Journal of Medicine*, *296*, 716–721.

Willan, A. R., & O'Brien, B. J. (1999). Sample size and power issues in estimating incremental cost-effectiveness ratios from clinical trial data. *Health Economics*, *8*, 203–211.

Williams, A. (2004). *What could be nicer than NICE?*. London: Office of Health Economics.

Williams, A., & Cookson, R. (2000). Equity in Health. In A. Culyer, & J. Newhouse (Eds.), *Handbook on health economics*, Vol. 1B (pp. 1862–1910). Amsterdam: North-Holland.

Winkelmayer, W. W. C., Weinstein, M. C., Mittleman, M. A., Glynn, R. J., & Pliskin, J. S. (2002). Health economic evaluations: The special case of end stage renal disease treatment. *Medical Decision Making*, *22*, 417–430.

Zinn, C. (2002). Plan to cut drug spending attacked by doctors. *British Medical Journal*, *324*, 937.