The (Near) Equivalence of Cost-Effectiveness and Cost-Benefit Analyses
Fact or Fallacy?

Cam Donaldson
Health Economics Research Unit, University of Aberdeen, Aberdeen, Scotland

Summary

There has been much recent debate in the health economics literature as to the (near) equivalence of cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA). The aim of this paper is to demonstrate that whether such a (near) equivalence exists depends on whether one defines economic evaluations as ‘CBA’ or ‘CEA’ on the basis of either what is measured or what question the analyst is seeking to answer. The former basis of definition is popular within the ‘decision science’ approach to economic evaluation, but does not seem to have any theoretical support. If the latter, more theoretically correct, basis is accepted, there is no longer a case for the (near) equivalence of CBA and CEA.

In the recent debate about guidelines for economic evaluation, the implication seems to be that there is much agreement about methods, and that either research efforts should concentrate on resolving remaining controversies (such as whether to include future medical costs arising from an intervention now) or such differences should just be more explicitly recognised as unresolvable.[1-3] However, such acceptance of a large amount of agreement about methods seems odd when, as reflected by the following quote by Phelps and Mushlin, health economists cannot agree on simple things such as whether there is any difference between cost-effectiveness (CE) and cost-benefit (CB) analyses: ‘We wish to lay to rest the fictitious separation between CE and CB analyses: for many purposes, they are equivalent.’[4]

The quote reflects the attitude of many in health economics that there is a (near) equivalence of cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA). It is the contention of this paper, however, that whether practitioners and commentators regard this equivalence as fact or fallacy depends on whether analysis types are defined in terms of: (i) what question is being addressed; or (ii) what is measured. Recently, as will be shown, there has been a trend in health economics to use the latter as a basis of definition. This seems to be driven by a practical ‘decision science’ approach to economic evaluation, i.e. an approach which seeks to reduce everything to 1 number (e.g. a cost-effectiveness ratio); a number which leads to a ‘decision’. Despite the apparent simplicity of this, it may also prove to be the undoing of economic evaluation, a point which I attempt to explain later in this paper.

It is the aim of this article to highlight how these 2 different sources of definitions of CEA and CBA can lead to the fallacy of their apparent (near) equivalence. This is done by, first, outlining definitions
of CBA and CEA according to what question is being asked and, second, demonstrating the conflict between such definitions and those based on what is measured. Finally, there is a discussion of why accurate definition of these techniques matters.

1. Definitions of CBA and CEA

CBA and CEA each have advantages and difficulties associated with them. But, as Mooney has said, ‘ease cannot be allowed to dictate use; it is a question of what is best for which question.’[5] It is the view of the author that, contrary to recent trends in health economics, which technique to use should be determined by the question to be addressed in an evaluation. This is consistent with economic theory.[6,7]

1.1 Cost-Benefit Analysis

CBA is used to determine allocative efficiency. It seeks to answer the question, ‘is it worth achieving this goal?’ or ‘how much more or how much less of society’s resources should be allocated to achieving this goal or to this type of healthcare?’[8] It involves interpersonal comparisons of preferences because allocation of more resources to 1 group implies a redistribution whereby that group gains but another group loses those resources.

In principle, the answers to such questions require all costs and benefits to be valued in a commensurate unit (such as money). The perspective is societal, since ‘... cost-benefit analysis is based on including all costs and benefits irrespective of to whom they accrue.’[8]

Thus, patient costs (e.g. travel and time off work) associated with treatment are included in CBA. If everything is measured in 1 unit, then in theory, comparisons are made easier. Conceptually, CBA has a very wide range of applicability as, with everything valued in commensurate terms, it could be used to compare healthcare objectives with each other or with those arising in other sectors of the economy.[9]

In practice, however, the monetary valuation of benefits in CBAs is difficult. There are problems with both the human capital and willingness-to-pay approaches.[10] Not surprisingly then, no ideal CBA has been carried out in the field of healthcare despite the titles of many articles bearing the name. In many ‘cost-benefit analyses’, the cost of a new intervention is often assessed against ‘benefits’ measured in terms of cost savings. However, this clearly involves only a comparison of costs. No consideration is given to the difficult issue of valuing health improvements and other benefits in monetary terms.[11]

Despite such problems, CBA remains a useful tool, particularly in setting out a decision-making problem. By identifying the costs and benefits associated with different healthcare programmes and valuing what can be valued, one can explicitly observe the trade-offs between costs and benefits (tangible or intangible) resulting from a decision to implement, or not implement, a healthcare programme. The following quote, while recognising the imperfections of CBA in practice, also highlights its importance as a framework for decision-making: ‘A good CBA will: identify relevant options for consideration; enumerate all costs and benefits to various relevant social groups; quantify as many as can be sensibly quantified; not assume the unquantified is unimportant; use discounting where relevant to derive present values; use sensitivity analysis to test the response of net benefits to changes in assumptions; and look at the distributive impact of the options.’[12]

A good example of this is the work on costs and benefits of introducing child-proofing of drug containers in the UK in the early 1970s. In 1971, a decision was taken not to introduce such child-proofing on the basis of its cost [500 000 pounds sterling (£) per year at the time]. Using conservative estimates, Gould[13] showed that without child-proofing there would be 16 000 hospital admissions per year at a cost of £30 per admission, resulting in costs of £480 000 per year. The extra cost of child-proofing was £20 000 per year. If only 20 lives could be saved by child-proofing, the decision not to
child-proof would imply the life of a child is valued at less than £1000.

The figure of £1000 would appear to result from what would now conventionally be called a ‘cost-effectiveness ratio’. But this ratio has little to do with cost effectiveness as defined in section 1.2. Because of opportunity cost, the extra resources required for child-proofing (i.e. the extra £20 000) would have to be found by depriving some other group(s) of patients of their (hopefully beneficial) care. What the analysis highlights is that one group loses out while another gains. Even though a monetary value was not placed on the life of a child, such trade-offs would almost certainly have remained implicit in the absence of CBA, thereby not allowing them to be subjected to the same degree of scrutiny. In the absence of such scrutiny, important aspects of efficiency may not be recognised. CBA, by analysing who receives benefits and bears costs, also opens up important issues of the distribution (or equity) of health and healthcare which would have remained uncovered. By some quantification and making values explicit, CBA can be a very useful decision-making aid, even without having everything valued in monetary terms.

It has to be acknowledged, of course, that if attempts are made to value benefits in monetary terms by the use of willingness to pay, an aggregate measure of benefit based on such values will, to an extent, reflect the current distribution of income in society. However, there is little investigation of whether measures of benefit used in other types of economic evaluations also reflect the distribution of income and, of course, analysts seem content to use cost data in all types of economic evaluations which are based on prices which reflect the income distribution!

1.2 Cost-Effectiveness Analysis

Cost effectiveness has a very specific meaning in economics. It is about identifying the point on an isoquant curve (a curve which traces out the combinations of 2 or more inputs that give rise to the same level of output) which is tangential to an isocost line (which shows the combinations of any 2 or more inputs that can be bought with the same amount of money). This amounts to minimising the cost of a given output (or goal). CEA then seeks to answer the question, ‘given that it has been decided that a goal is to be achieved, what is the least-cost way of doing so?’. Cost effectiveness implies technical efficiency which is about ensuring the production of the same output with less of one input and no more of others. Cost effectiveness and technical efficiency each imply output efficiency which is about ensuring that the output is the maximum which a combination of inputs is capable of producing.[14] This leads to questions similar to those addressed in CEA, and often addressed under the guise of CEA, such as, ‘what is the best way of spending a given budget?’.

CEA always involves comparison of at least 2 options with the same goal (or the same budget). With respect to perspective, there is strictly no role for patients’ costs in CEA. Patient costs come from a budget which is different to that of the care being evaluated. In practice, it is difficult to argue against the inclusion of patient costs, although it is not clear how such information should be used.

Patient costs really become relevant when results based on their inclusion conflict with the results of CEA (which in theory would exclude them). The policy question then becomes one of whether effects on patient costs carry more weight than the gains to be had without considering such costs. This will often turn out to be a CBA question, which makes recent recommendations to include such costs in CEA rather odd.[15] For example, let us say that a study has shown day hospital care for elderly people to be equally effective but less costly (i.e. more cost effective) than inpatient care, when considering healthcare costs only. However, let us also say that it has been proposed not to have such day hospital care for elderly people on the grounds that the patient (and carer) costs of maintaining the patient at home are too great. The proposal would lead to greater provision of more costly inpatient beds than would have otherwise
been the case. But this means that, because of the extra resources now tied up in inpatient care than would otherwise have been the case, another group of patients somewhere in the healthcare system would lose out, as these resources would have some other use. The question then becomes one of whether the proposal is worthwhile (a CBA question) rather than one about the best way of providing care for elderly people within a given budget (a CEA question).

CEA can take 2 main forms. In the first, if the health outcomes of the alternatives to be compared are known to be equivalent, cost differences only need be analysed. This is sometimes known as cost-minimisation analysis. The least costly alternative is obviously most efficient as resources are saved which can be put to some other beneficial use without reducing the health outcomes of the client group being studied. In the second form of CEA, alternatives may differ in terms of cost and effect. A ratio is produced for each alternative, in which the numerator is cost and the denominator is the health effect under consideration. Health effects are measures of final outcome: most commonly, they are quite narrow measures such as ‘life-years saved’, ‘heart attacks prevented’ or ‘improved physical functioning’. The ‘cost-effectiveness ratio’ produced for each alternative is therefore a measure of ‘cost per unit of health effect’. The alternative with the lowest cost-effectiveness ratio is best. Within a given budget, greater health can be achieved by implementing this alternative (thus enhancing output-efficiency as well as cost-effectiveness). For example, a common early use of a ‘cost per life-year saved’ ratio was in comparing alternatives for the treatment of chronic renal failure. Of course there is, at this stage, the possibility of reducing the size of the budget, producing the same level of output and allocating the resources to another programme. This demonstrates the inextricable link between cost effectiveness, technical efficiency, output efficiency and allocative efficiency (but it does not mean they are the same thing!). When outcomes of alternatives, even within a budget, have multidimensional effects, an index [such as quality-adjusted life-years (QALY), healthy years equivalents (HYE) or willingness to pay (WTP)] where values are elicited from patients (or prospective patients) could be used to aid judgements about the most cost-effective ways of spending the budget. Thus, although it may seem contradictory at first sight, techniques such as willingness to pay can be used within a CEA framework.

The most important point, however, is that when using CEA, no interpersonal comparisons of utility are made. Of course, it may be thought that in practice such comparisons cannot be avoided. In CEA, the assumption that the unit of outcome means the same to each individual does involve interpersonal comparisons; that is, when the information is used within a cost-benefit framework. If the analysis is about demonstration of the least-cost way of achieving a given outcome or maximising a given outcome from a given budget (as in CEA), the differing values attached to such an outcome by different individuals is unlikely to matter, within the limitations of the analysis, as it is the same group who will be treated anyway. No one will be denied treatment in order to benefit some other group of patients. Surprisingly, this is an insight which is overlooked in the influential work on CEA by Gold et al. Once cost effectiveness and output efficiency are achieved, we can only move to a higher isouquant (i.e. a higher level of production or outcome) for a programme by allocating more resources to it. This does involve interpersonal comparisons as those resources could have been used in an alternative programme. This brings us back to comparing groups again and only the CBA framework can address such comparisons.

1 Or a slightly expanded group of similar patients if the CEA is about doing more with the same costs. Of course, as pointed out in the previous paragraph, if outcomes are multiple, a QALY, HYE or WTP technique may be used (which will imply interpersonal comparisons).
2. Definitions According To What Is Measured

The primary distinction between the 3 socio-economic techniques lies in the valuation of the health-related consequences of the intervention under study. In CEA, consequences are expressed in natural health units (e.g. years added to life expectancy); in cost-utility analysis (CUA), consequences are expressed in patient preference measures combining length and quality of life (e.g. QALY); and in CBA, all consequences are expressed in monetary terms.²⁰

The above quotation is an example of defining techniques of economic evaluation by what is measured. This practice is popular within health economics and is reflected in several guidelines for economic evaluation.²¹,²²

How does basing definitions of CBA and CEA on what is measured conflict with basing such definitions on the question addressed? In the description of CEA in section 1.2, the phrase ‘within a given budget’ is of crucial importance as is the fact that (in theory) CEA is about whether a given group of patients can be treated more efficiently. Often, authors produce a ratio of extra costs per extra unit of health effect for one intervention over another, calling their analysis an ‘incremental CEA’. An example of this is given in the case of different strategies for reducing serum cholesterol levels in table I.²³ A population intervention produces an extra cost per extra year gained (over the status quo) of £12 and an extra cost per extra QALY gained of £10. The dietary intervention produces an extra cost per extra life-year gained of £12 440 over and above the population intervention and an extra cost per extra QALY gained of £100 546 over the population intervention. However, although the data in table I are referred to as ‘cost-effectiveness ratios’ they are not really so. All of these interventions would require more resources to be allocated to preventing heart disease relative to the status quo. This requires a movement in isquants which will mean that the resources to meet these extra costs will inevitably come from some other healthcare programme (i.e. another budget). Some judgement is required as to whether such extra costs are worth incurring. This takes us back to broader (interpersonal) comparisons, which is the role of CBA, not CEA. The distinction does not necessarily change what data are collected in economic evaluations conducted alongside randomised trials. Rather, it is a case of recognising that such data have to be used within a broader framework which may require nontrial- (or different trial-) based information on alternative uses of resources and equity considerations. It is more accurate to say that data on incremental costs and outcomes will be used in a cost-benefit framework to aid a decision about whether such incremental costs are worth incurring.

Perhaps this also helps to answer the question raised recently as to why interventions with relatively low incremental cost-effectiveness ratios are not expanded in scope at the expense of more costly interventions.²¹ It is not just a case of establishing a low cost-effectiveness ratio and saying that programmes with such ratios should be implemented/expanded. The decision is more difficult, in that faced with a fixed budget, the decision-maker is going to have to take resources from some other programme to fund the one with the low ratio. This takes us back to allocative questions which are more difficult to address than is recognised by a simplistic ‘cost-effectiveness framework’.

3. Semantics or Substance?

Defining techniques of economic evaluation according to what is measured gives the impression

<table>
<thead>
<tr>
<th>Table I. Incremental cost effectiveness and cost utility of programmes aimed at reducing serum cholesterol levels²³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>No intervention</td>
</tr>
<tr>
<td>Population intervention</td>
</tr>
<tr>
<td>Dietary advice</td>
</tr>
<tr>
<td>Dietary advice plus drugs</td>
</tr>
</tbody>
</table>

Abbreviations: QALY = quality-adjusted life-year; £ = pounds sterling (year of costing 1990).
that CBA is rarely possible, when in fact, analysts are carrying out studies using the framework of CBA all the time. They just call them ‘incremental CEAs’. This is not, in itself, problematic and does indeed boil down to semantics if it makes no difference to decisions and if it does not harm perceptions of the usefulness of economic evaluations amongst policy-makers.

However, perceptions may be harmed in 3 ways. First, decision-makers are (understandably) led to believe that if an intervention has a low cost-effectiveness ratio then it should be implemented. Looking at different ways of treating the same group of patients, it is obvious that, other things being equal, the most cost-effective intervention should be implemented. It can simply be substituted for less cost-effective care. Results of studies in which cost and effectiveness are incremental to the status quo cannot be implemented so easily. Even if an intervention has a low incremental cost-effectiveness ratio, extra resources must be found from somewhere else. Examples of this are recent studies in which incremental cost-effectiveness ratios are produced for simvastatin over no intervention[24] and alteplase (tissue plasminogen activator; t-PA) over streptokinase[25] for patients postmyocardial infarction along with the following statements: ‘... treating patients with coronary heart disease with simvastatin is cost-effective in both men and women at the ages and cholesterol levels studied’[24] and that alteplase over streptokinase, ‘... would be cost effective by customary criteria.’[25]

In the former case, resources of just over $US1500 per patient (1995 values) would be required to implement the programme, while in the second case, the authors themselves state that, for the US, implementation of an alteplase programme, ‘would cost the nation approximately $500 million each year’.[25] Thus, something will have to sacrificed, like a drug which is less effective for the resources spent (perhaps those which are prescribed for mild hypertension) or some other nonpharmaceutical treatment (like some types of minor surgery for some groups of patients) in order to fund these ‘cost-effective’ interventions (unless taxes or insurance premiums are to rise which implies an opportunity cost of a different sort, but still an opportunity cost). Decision-makers and other ‘customers’ for economic research should be informed of this at the outset of such studies. It needs to be made clear that such information, on incremental costs and incremental effectiveness, can be used only to aid decisions about whether it is worth allocating more resources to the programme of care within which the intervention evaluated falls. Cramming results into mechanistic cost-effectiveness ratios and, even worse, into ‘league tables’,[26] will not avoid the need to make such decisions. Sometimes economics can do no more, and this realisation may lead to less disappointment with results from economic studies.

A good example of the above problem is in the area of incorporating paclitaxel into first-line combination therapy for advanced ovarian cancer. Two recent studies conclude, on the basis of estimated incremental cost-effectiveness ratios, that this therapy is cost effective.[27,28] A third study has shown that, for the Canadian province of Ontario alone, use of this therapy would cost 9 million Canadian dollars ($Can) per year (1993 values) over and above the present costs for this group of patients.[29] From where will the resources come? The authors of this third study[29] propose that other areas of the chemotherapy budget be examined first and, if care is thought to be optimal within that budget, the budget for treating all aspects of ovarian cancer should then be considered as a source of sacrifices in order to fund paclitaxel and so on through the budgets for gynaecological cancer, all cancers and the total healthcare budget. Other authors have shown that if therapies, which have been claimed by the authors of studies to be ‘cost effective’, are simply implemented without consideration of the opportunity costs, the likely result is simply an escalation of the total healthcare budget without due consideration being given to whether this is desirable.[30,31]
One further problem is that to say CBA involves all consequences being valued in terms of money can encourage, amongst the uninstructed, more of the limited types of CBAs referred to previously, whereby the costs of an intervention are measured against the costs saved (defined as ‘benefits’) from undertaking less of some alternative intervention. No attempt is made to assess welfare effects or ‘intangibles’ and, often, they are never mentioned.

There is also the danger that many economic evaluations will be misclassified in registers of such evaluations currently being compiled by the Centre for Reviews and Dissemination at the University of York and the Office of Health Economics in London, UK. Some studies may prove difficult to classify at all if done so on the basis of what is measured. This classification may be made easier to do by doing so on the basis of the question addressed. Indeed, it may be more useful to give some studies a multiple classification as the information produced in them could contribute to addressing cost-benefit, as well as cost-effectiveness, questions.

Overall, if one’s view is that economic evaluations are defined by what is measured, there is a strong case for the (near) equivalence of CEA and CBA. However, if one’s view is that evaluations are defined by the question the analyst is seeking to answer, the case no longer holds. If there is a theoretical basis for the latter definition, then the (near) equivalence argument is based on a fallacy.

Acknowledgements

The Health Economics Research Unit (HERU) is funded by the Chief Scientist Office of the Scottish Office Department of Health (SODoH). However, the views expressed in this paper are those of the author, not SODoH. The author is grateful to colleagues at HERU for comments received at an internal seminar some time ago. The paper was revised while the author was on sabbatical at the Centre for Health Economics.

2 In practice, this may not be a problem. If people search such databases by clinical problem, they are likely to take all studies (whether CEAs or CBAs) in that area and use them as they see fit. This is because, for any given clinical area, there are likely to be few economic studies. However, as the numbers of studies on these databases grow, so might the problem.