ON THE (NEAR) EQUVALENCE OF COST-EFFECTIVENESS AND COST-BENEFIT ANALYSES

Charles E. Phelps
Alvin I. Mushlin
University of Rochester School of Medicine and Dentistry

Abstract
Many people believe that cost-effectiveness (CE) and cost-benefit (CB) analyses require different assumptions. However, when CE analysis supports decisions to use medical resources, it makes the same assumptions that CB analysis requires. They are mathematically equivalent. Differences between CE and CB hinge more on reporting style than on fundamental assumptions.

Medical literature now contains numerous cost-effectiveness (CE) analyses, often with the explanation that cost-benefit (CB) studies have not been carried out because they require stronger assumptions. Yet the two forms of analysis are very similar and, in many cases, require the same assumptions to draw meaningful conclusions about the allocation of resources.

We illustrate this similarity by reference to medical decision analysis. The use of a CE ratio to select among competing medical interventions (or the "cutoff" CE ratio) makes the CE analysis equivalent to a CB analysis. To demonstrate this near equivalence, we develop the standard CE and CB models, show their uses in making resource allocation decisions, and compare their results and the assumptions needed to reach those results.

THE METHODOLOGICAL DILEMMA

The CE approach to decision-making is widely used on medical problems, but, surprisingly, formal definitions of the concepts underlying CE analysis have received little attention. In their widely used text, for example, Weinstein and Fineberg (8) stated:

The underlying premise of cost-effectiveness analysis in health problems is that, for any given level of resources available, the decision maker wishes to maximize the aggregate health benefits conferred to the population of concern. . . . Alternatively, a given health benefit goal may be set, the objective being to minimize the cost of achieving it. (p. 239)

This research was supported by the National Center for Health Services Research and Health Care Technology Assessment, Department of Health and Human Services, Grant No. 1-R01-5477.
Table 1. Adoption of a New Medical Intervention

<table>
<thead>
<tr>
<th>Benefits increase</th>
<th>Benefits decrease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costs increase</td>
<td>Conduct cost-effectiveness</td>
</tr>
<tr>
<td></td>
<td>analysis</td>
</tr>
<tr>
<td>Costs decrease</td>
<td>Adopt</td>
</tr>
<tr>
<td></td>
<td>Conduct cost-effectiveness</td>
</tr>
<tr>
<td></td>
<td>analysis</td>
</tr>
</tbody>
</table>

They go on to describe how utilities are measured in one unit (e.g., quality-adjusted life years [QALYs]) and cost in another. They then offered the following rule for making resource allocation decisions using CE analysis:

Alternative programs or services are ranked from the lowest value of this cost-per-effect ratio to the highest, and then they are selected starting with the highest ranked program or service until available resources are exhausted. *The point on the priority list at which the available resources are exhausted, or at which one is no longer willing to pay the price for the benefits achieved, becomes the cutoff level of permissible cost per unit of effectiveness.* (p. 239; emphasis added)

Usually a CE analysis assesses the incremental costs and incremental effectiveness of a new strategy, device, or program, relative to one previously used. Comparisons of the changes in benefits and costs lead to four common alternatives, as shown in Table 1.

This approach says to adopt a new intervention without further analysis if (and only if) benefits increase and costs decrease. In this case, the new intervention *dominates* the old. Conversely, if costs are higher and benefits lower with the new intervention, it is dominated by its predecessor and should not be used.

CE analysis requires further information when costs and benefits both rise or when they both fall. When both benefits and costs rise, the analyst must calculate costs and effectiveness to see if the new intervention has an incremental CE ratio at least as favorable as others competing for the health care budget (or matching some preselected cutoff). Similarly, if both costs and benefits fall with the new intervention, this should prompt (if not previously undertaken) an incremental CE analysis of the *old* intervention (relative to the new) to see if it matches desired standards of cost-effectiveness. The same decision rule then follows in both cases: adopt the strategy so long as the incremental CE ratio does not exceed the desired standard of comparison (or the budget is not exhausted).

In CB analysis, benefits as well as costs are measured in dollars and, thus, can be added and subtracted with no use of ratios. As Weinstein and Fineberg (8) stated: “Once benefits and costs have been expressed in monetary terms, net benefits are derived as the difference between the two: If these are positive, the argument goes, the program or practice should be undertaken; if they are negative, it should not.” They continue this discussion with a statement that we believe is widely shared in medical circles: “The major disadvantage of the benefit–cost framework is the requirement that human lives and the quality of life be valued in monetary units. Many decision-makers find this difficult or unethical and do not trust analyses that depend upon such valuations” (pp. 239–240). Similarly, Warner and Luce (7), when discussing CB and CE analysis, stated, “If placing a value on [health] benefits is regarded as impossible or undesirable, CB [analysis] can be rejected and CE [analysis] adopted” (p. 87).
The remainder of this article discusses the basis for this mistrust and examines whether in fact more similarities than fundamental differences exist between CE and CB analysis.

A FORMAL STATEMENT OF CE ANALYSIS

A formal characterization of CE analysis reveals its similarities to CB analysis more fully. This analysis usually compares a “competing” technology or intervention with an established one, which we call e (established) and n (new). We call the well-being (utility) produced by the two interventions $U_n$ and $U_e$, respectively, with respective costs $C_n$ and $C_e$. The usual decision rule says to proceed with the new approach if the ratio of added benefits (from using a new intervention), divided by the added costs, exceeds a prespecified level, that is:

$$\frac{U_n - U_e}{C_n - C_e} \geq g$$

(1)

where $g$ is the “cutoff” ratio for effectiveness/cost that is “acceptable.” This is the inverse of the ratio expressed in Weinstein and Fineberg (8) but otherwise represents the same concept. Typically (and we believe, correctly), $g$ represents the effectiveness/cost ratio (ECR) of the least desirable intervention currently undertaken. That is, $g$ represents the ECR of the first activity that would be reduced or eliminated if resources declined. It represents the cutoff level of effectiveness per unit of permissible (or budgeted) cost.

We now slightly reformulate equation (1), recalling that the CE “prescription” says to proceed if the inequality in equation (1) holds. Now multiply both sides of the equation by $(C_n - C_e)$, giving:

$$U_n - U_e \geq (C_n - C_e)g$$

(2a)

then rearrange terms to give:

$$U_n - C_ng \geq U_e - C_eg$$

(2b)

Equation (2b) says to use the new intervention (n) if its “net benefits” exceed those of the established technology (e). The concept of “net benefits” explicitly recognizes that the resources used to produce the benefit come at the “cost” of some utility that those resources might have produced with other uses. The “cutoff” value $g$ represents the ECR of those alternatives—the “opportunity cost” of resources. Thus, $C_ng$ represents the foregone benefits from another activity if resources $C_n$ are used in the new activity $n$ rather than in the other.

We can restate these ideas in more familiar CE terminology. Begin with the usual CE prescription to proceed if

$$\frac{\text{Incremental Effectiveness}}{\text{Incremental Cost}} \geq g$$

(3a)

which is equivalent to

$$\text{Incremental Effectiveness} \geq (\text{Incremental Cost}) \times g$$

(3b)
Near equivalence of CE and CB analyses

which in turn corresponds to

\[
\text{Incremental Effectiveness} - (\text{Incremental Cost}) \times g \geq 0 \quad (3c)
\]

or

\[
\text{Net Incremental Effectiveness} \geq 0 \quad (3d)
\]

This formulation of the CE approach shows that it is equivalent to the CB approach, where benefits are added and subtracted, \textit{once a specific value of g is selected}. Apparently, this step makes some analysts using CE methods uncomfortable. Yet, one cannot use CE methods to make decisions about resource allocation without selecting a specific value of g.

These formulations are based on \textit{incremental} cost and outcomes, \textit{measured as "small" changes} in the way resources are allocated. For decisions involving large changes in the pattern of resource use, the "marginal" value for g might not be appropriate but rather some higher value for g. The "acceptable value" of g might be based on a ratio averaged across several (increasingly valuable) alternative uses of resources. This simply recognizes that any large change adds up a series of smaller changes, each valued appropriately.

**DECISION-MAKING USING CE ANALYSIS**

We next discuss the use of decision-making based on CE methods in several different situations.

**Situation 1: Constant Incremental Benefits across Population.** Suppose two diagnostic tests are available with constant expected benefits no matter what fraction of an eligible population receives the tests. The CE approach says to apply the "best" test (on the basis of incremental CE ratios) until all the population is served or the budget runs out.

**Situation 2: Procedures with Declining Benefits.** Consider now a case where the incremental benefit declines as the number of treatments expands. Suppose two interventions are available, both with declining effectiveness as use expands. For example, the use of a 6th stool guaiac produces fewer lives saved through cancer detection than the 5th, and similarly for the 4th, etc. (4). Mammography finds decreasing amounts of breast cancer as the age criterion for application expands to include younger women (1,2). Appropriate use of these alternative methods to produce QALYs would first begin with the procedure with the best ECR ratio. Its use would expand until the ratio of incremental effectiveness to cost just matched that of the other intervention. At this point, additional resources would be used to increase the amount of both interventions. Expanding each so that the ratio of incremental effectiveness/incremental cost remained the same for each possible intervention.

The case highlights a particular phenomenon about CE analysis: \textit{if} the scope or intensity of an activity can be adjusted continuously (i.e., decisions about the extent of its use are not lumpy) and \textit{if} the activity has declining marginal effectiveness, or increasing marginal cost, or both, as the scope or intensity expands, \textit{then} the appropriate adjustment of these activities leads them \textit{all} to have the same incremental ECR. The investment in every activity with these characteristics should be adjusted so that the next application of it is "marginal" in the sense that its ECR = g. This approach would maximize net benefits to the population under consideration.
Quasi-Factual Example Using Cancer Detection Methods

To make this discussion more concrete, we now offer a specific example of CE-based decision analysis. This example relies on data that are “approximately” accurate, but it is only intended to provide a concrete example of the more general concepts discussed earlier, rather than a careful analysis of the alternatives discussed. Suppose a cancer detection program was considering expanding the use of Pap smears but (because of budget limitations) would have to reduce the use of stool guaiac tests to do more Pap smears. According to one study (adjusted for inflation), the cost per QALY from the 5th stool guaiac is about $500,000 (4). Suppose we eliminate such testing in a hypothetical population, saving $1 million. We also expect (in the statistical sense) to lose 2 QALYs in the process.

We now take this $1 million in “saved” money and apply it to Pap smears. By recent estimates (2), the incremental cost per QALY with Pap smears is about $20,000, a figure that depends, of course, on the screening strategy selected. With this alternative use of resources, we can gain 50 QALYs with $1 million worth of Pap smears. Note that, if we further expanded the use of Pap smears, the cost per QALY might well increase, as we extended the use of the test to a less susceptible population or increased its frequency in a given population.

The net benefit from this reallocation of resources is 48 QALYS, that is, the 50 QALYs gained less the 2 QALYs “opportunity cost.” The CE decision rule would tell us to proceed with this switch, since more QALYs can be achieved for the target population with the same budget by making this switch. While we have not conducted a formal analysis of this decision, it should be obvious that the same examination could be conducted for funds on the 4th stool guaiac, then the 3rd, etc. At some point, stool guaiac tests will likely become at least as incrementally cost-effective as Pap smears at producing QALYs, since the incremental effectiveness of the stool guaiac test rises as the intensity of use falls, and conversely for the use of Pap smears with an expanding population and rate of use.

DECISION-MAKING USING CB ANALYSIS

A CB analyst would typically proceed down a different path to analyze a medical intervention, but also typically will get to the same point as a CE analyst, so long as both agree on the marginal value of QALYS. CB rules say to adopt a medical intervention if the net benefits exceed the net costs. CB analysts commonly calculate the dollar value of benefits by learning the amount of benefits (e.g., QALYS) and then valuing them at some predetermined rate, such as $25,000 or $50,000 per QALY. For example, Warner and Luce (7) stated in their text on CB applications in medical care:

The measurement of many types of benefits is straightforward. Tangible personal health benefits can be measured in terms of years of life saved, months of disability avoided, or days of morbidity averted. Theoretically, health care resource benefits should be measured in physical units reflecting the resources saved, such as physician time, days of hospital bed occupancy, and so on. . . . Other economic benefits are counted in terms of obvious units; in the case of productivity losses averted, days of work loss avoided serves as a useful measure. In the case of intermediate outcomes, measurement again relies on obvious units, such as the number of accurate tests, in the case of diagnostic techniques, or pounds lost, in the case of a weight control program. (p. 85)

Several approaches exist for assigning a particular value for each unit of economic, biological, or clinical benefit. This value may come from an estimate of the earnings
potential of an individual (the so-called human capital approach), or from other methods. The human capital approach is generally viewed as an incomplete determination of value of health-producing activity, and most analysts now turn to methods such as willingness-to-pay (WTP) questionnaires, or indirectly estimated WTP values.

As Warner and Luce (7) noted, the assignment of value for benefits is difficult, but forms a central part of the analysis:

In CE [analysis], analysts do not attempt to value program effectiveness in monetary units. Consequently the task of effectiveness assessment ceases when the measurement phase is completed. In CB [analysis], by contrast, translating outcome quantities into dollar values constitutes a central feature of the analysis. (p. 86)

The calculation of a medical intervention's costs proceeds similarly, with the units of inputs measured or estimated, and then valued at their market cost or opportunity cost. Typically, the valuation of costs is less controversial than benefit evaluation, since markets exist for most inputs. The only important problem emerges when joint or fixed costs appear, in which case the analysis must account for these costs appropriately. However, these problems are identical in both CB and CE methods.

CB methods—like CE methods discussed earlier—require a somewhat different treatment when large (rather than marginal) changes in resource allocation are considered. In the usual CB framework, total benefits and total costs are compared, which requires adding up the incremental values for different quantities. In the pure economics framework, this can be accomplished by estimating a demand curve and then finding the total value of the medical intervention in question by deriving the area under that demand curve (3). The same approach remains valid if a WTP method is used, so long as the WTP measure can be made to vary with total quantity. In the most complete analyses, the WTP measure depends on the total amount of the activity or service provided, so that both small changes (using marginal WTP) and large changes (using the area under the WTP curve) can be evaluated accurately. Thus, like our previous discussion of CE methods, we need to be careful to distinguish between small "marginal" changes (where the incremental value of the good or service can be assumed to remain constant) and large changes (where the value per unit will decline as the amount consumed expands). However, nothing differs between CB and CE analysis in this respect.

THE NEAR EQUIVALENCE OF CB AND CE METHODS

We hope that the similarities between CB and CE methods have become more clear through the previous discussions. Here, we draw the comparison more closely in several specific cases.

**Dominance Cases.** In the most simple situation where CE methods provide clear-cut decisions — the case of dominance — CB methods provide the same answer, *no matter what cutoff value g has been selected*. To see this most easily, look at equation (2a). There, if incremental benefits are positive ($U_b > U_c$) and incremental costs are negative ($C_b < C_c$), the case of dominance, then equation (2a) always says to proceed, no matter what value of $g$ is selected, since the right-hand side of equation (2a) is always negative, and the left-hand side is always positive. For any value of $g$, when dominance occurs, the net benefits of activity $n$ exceed the net benefits of activity $e$, which is equivalent to saying that a CB analysis of changing from $e$ and $n$ would say to proceed.
Cases Without Dominance. As Weinstein and Fineberg noted, many decision makers are queasy about assigning dollar values to lives and, hence, distrust analyses that depend upon CB criteria. We view this as a misplaced concern, since the same assumptions are needed to use CE analysis to choose among competing health interventions. Using any CE analysis directly for resource allocation requires the same value judgments as those to make decisions using CB analysis.

In general, the major difference between CE and CB methods is that CB analysis states the dollar value of a life (or QALY) explicitly, whereas CE analysis does so implicitly. Or, put slightly differently, analysts use CE methods to estimate "prices" of QALYs from different activities. Decisions in the CB framework depend upon net value, whereas CE analysis asks if the price is acceptable. We can show that these are equivalent in any of several settings where CE methods are commonly used.

Suppose a new technology (or therapy) is being considered or an old one might be expanded. CE methods say to proceed if the ratio of marginal effectiveness to marginal cost is "as good as" other commonly used technologies or therapies. Weinstein and Stason (9) provided a summary of this standard model:

The criterion for cost-effectiveness is the ratio of the net increase of health-care costs to the net effectiveness in terms of enhanced life expectancy and higher quality of life. The lower the value of this ratio, the higher the priority in terms of maximizing benefits derived from a given health expenditure. (p. 718; note that they use a ratio of costs to benefits, while we discuss a ratio of benefits to costs)

This approach clearly compares CE ratios across different medical activities. The "typical" CE ratio for other medical activities helps estimate the cutoff level g. If no budget has been set formally, the ranking of projects suggested by Weinstein and Stason serves the same function as a cutoff, by providing a target level g for any project: it should be as good as all remaining competitors for resources. But again, each CE ratio selected implies a dollar value per QALY and, thus, provides the necessary information to conduct a straightforward CB analysis.

Next, it is most clear from our analysis that when a pure budget constraint is set (as in the VA or military hospital systems, the British NHS, an HMO, or even by a DRG payment for hospital services within Medicare Part A), then an implicit value of QALYs is set by the CE ratio assigned to the marginal project, the last one funded under rational decision-making, given the budget constraint. If medical activities are chosen in appropriate sequence and a fixed budget exists, then the "last" desirable project defines a cutoff CE ratio g.

In either case, resource allocation using CE methods requires or produces a value of g that would in turn allow direct use of CB methods. Thus, CE and CB methods are indistinguishable when used for selecting among competing health interventions in a resource-scarce environment.

Do other uses of CE methods exist? We think not, in a meaningful sense. Either CE studies are used to guide the allocation of resources, or they stand alone without such purpose.

HOW DO CE AND CB METHODS DIFFER?

Reporting Style
Despite the similarities between CE and CB (many of which have been recognized elsewhere), many people still have great difficulty with the idea that the two approaches
are “the same.” We have noted one difference previously: CB analysis typically determines “in advance” the marginal value of a QALY or a life, or other form of benefit, and then calculates the net benefits.

Alternatively, CE analysis typically calculates the “price” of a QALY or life and leaves the decision unstated. For this reason, CB statements depend on the value of QALY (or life) selected. As a reporting device, CB analysis typically adopts one (or a few) estimates of the appropriate value of a QALY and then calculates the net benefits. (For large-change projects, one WTP curve can be chosen, from which total benefits are calculated. The analog to changing “the value” of a QALY in such a case is to shift the WTP curve up or down in a sensitivity analysis.) By contrast, CE analysis usually leaves the statement in the form of a price, that is, dollars per QALY from some particular activity. The decision maker must then decide whether this is “too high a price” in allocating resources. However, as we have shown, as soon as decisions are made on the basis of the CE analysis, the same value judgments have been made as if a CB analysis had been conducted. The common difference in methods of reporting CE and CB results does not represent any true analytic difference. We believe that the common practice of conducting sensitivity analysis in CE studies would form a fruitful addition to the style of CB analysts, particularly across the valuation of benefits, where controversy exists, such as in the valuation of QALYs. But this is merely a difference in reporting style, not philosophy.

**Aggregation**

A second prevailing difference is not intrinsic to the methods of CB and CE but nevertheless appears in common practice: the level of aggregation. In concept, both CE and CB can be applied at either the level of the individual, the society, or anything in between. Nothing distinguishes one from the other in terms of the types of analysis that can be conducted.

However, many practitioners of CE analysis carry out their analysis at a highly disaggregated level, often for an individual patient, while CB analysis often applies to large groups, often “society.” The act of aggregating individuals to a societal level requires making interpersonal value judgments. In CB methods, a typical approach treats the next dollar of benefits to each person as the same, no matter who is the recipient (3). Others object to this particular method of aggregation but would approve aggregation with different “weights” for each person. For example, the weights might provide larger incremental benefit to persons with lower incomes or lower health status. The philosophy of Rawls (5) emphasized such approaches, stating that changes in policy should benefit the worst-off members of society first.

Still others object to any aggregation, at least formally, and would carry out CE analysis only for individual persons. It may be that the most common concerns about CB analysis arise because CB studies typically focus on large populations and, thus, have aggregated people “inappropriately” by the standards of the reader of the study. However, this is a separate question than whether resource allocation choices based on CE methods differ from those based on CB methods.

The most common application of CE methods appears in medical decision theory. Indeed, the CE ratio test (and the questions of dominance) are standard features of computer software designed to implement medical decision theory. These tools can obviously be applied to the individual patient. However, they can—and commonly are—applied to larger populations as well, and decisions regarding the uses of technologies commonly rest on comparisons with the CE ratio of “other” medical interventions, which in turn implies cross-person comparison. For example, the comparison
of cost-effectiveness of Pap smears for cervical cancer with pneumococcal vaccine for the elderly or polio vaccine for children obviously requires interpersonal comparisons just like aggregated CB analysis makes. Thus, CE methods applied to resource allocation decisions commonly involve aggregation just as much as CB methods.

**Multiple Dimensioned Benefits**

A third difference between CB and CE methods may persist in some settings. If a medical intervention produces gains in more than one dimension, CB methods provide a natural way to make them all commensurate, by converting each type of benefit (extensions in life expectancy, reduction in sick loss days, and even reduction of uncertainty about a diagnosis) into a common metric—usually dollars. CE analysis has a more difficult time with some of such conversions, but the goal of making all health outcomes commensurate clearly exists. The considerable literature providing methods of “quality adjustment” to calculate QALYs provides relative values within the sphere of health outcomes (6). These comparisons, usually derived by “standard gamble” approaches, elicit (for example) the relative weight people place on a day in bed versus a healthy day, and with those weights (to continue the example), assign .7 QALYs for each year spent in bed versus 1 QALY for a healthy year. To the extent that such adjustments to different components of health gains can be made, then all health gains can be converted to QALYs, and the incremental CE ratio g provides the final conversion to dollars (or in the reverse direction, dollars to QALYs).

**Joint Costs**

Finally, in the case of joint production (gains in more than one dimension of health, or even in health and something completely different), CE methods are not well equipped to deal with the analysis. The correct CB approach would add up all of the benefits from all dimensions of a project and compare those against the total costs. This task is easily accomplished when all benefits and costs have the same unit of measurement, for example, dollars. CE analysis would typically offer the marginal CE ratios along the several dimensions but not offer a direct recommendation. The difficulty in CE analysis comes from deciding what “incremental costs” really are when joint production exists. Typically, the analyst makes arbitrary assignments to divide the costs up among the various areas where benefits are produced, but any such allocations are precisely that — arbitrary — and have no particular meaning. CB analysis, by comparing total benefit and cost of an entire project, avoids this problem.

**CONCLUDING COMMENTS**

We wish to lay to rest the fictitious separation between CE and CB analyses: for many purposes, they are equivalent. We recognize that different analysts use these tools differently, but our descriptions of each approach capture the fundamental concepts of CB and CE analysis. As these comparisons show, decisions made about medical resources using CE analysis are wholly analogous to those using CB analysis. Neither is more or less value-laden, and neither intrinsically requires different levels of aggregation; rather, they express the assumptions and the results differently, but it all represents the same information. The literature of medical decision theory and technology assessment would be well served if it set aside the artificial dispute about using CE or CB methods and turned to the harder problems of stimulating and conducting such studies.
REFERENCES