The Limits of Cost-Effectiveness Analysis

William S. Weintraub, MD; David J. Cohen, MD, MS

The competitive worldwide economic environment and ever-increasing costs of health care have created a setting in which understanding costs and making sure that we achieve good value in health care are paramount. One approach to seeking value is through the use of cost-effectiveness analysis. Although this science is now several decades old, it has been refined over the last several years, with increasingly sophisticated statistical and standardized methods. Is cost-effectiveness analysis useful? Does it help in medical decision making and in allocation of scarce resources? In the accompanying article, “Cost, Effectiveness, and Cost-Effectiveness” Diamond and Kaul argue that cost-effectiveness analysis is not a useful approach. Although we agree with many of the points that Diamond and Kaul raise, we do not agree with their conclusion.

Cost-effectiveness analysis involves an assessment of both cost and effectiveness. The distribution of each needs to be understood. A cost-effectiveness analysis is only as valid as its underlying measures of effectiveness and cost, a discussion that is beyond the scope of this article. However, the methods to make these assessments vary considerably. There are standards for cost-effectiveness, but at times, perfectly adhering to these standards is not realistic, and compromises are often made that may be entirely scientifically legitimate.

Cost-effectiveness is, by nature, incremental. Thus, it is necessary to look at the added costs compared with a control group. Selection of the appropriate control group is a challenge itself. At times, the appropriate control is placebo, and at other times, it is active therapy; the appropriate control is dependent on the clinical question being asked. However, when cost-effectiveness analysis is conducted using data from a clinical trial, the selection of the control group will not be a decision that the analyst can affect (at least after the trial has been completed).

When additional costs and incremental measures of effectiveness of a new form of therapy are available, along with description of the distribution of each, an incremental cost-effectiveness ratio (ICER) may be measured (along with its own distribution). An ICER is a ratio of the incremental cost of the new therapy divided by the incremental measure of benefit. When the measure of benefit is expressed in life-years or quality-adjusted life-years, the ICER will be measured in cost per life-year or quality-adjusted life-year gained.

The ICER should not be viewed only as a single number, as there is uncertainty to both the measure of the cost and the measure of effectiveness. The first level of uncertainty is based on chance or sampling error alone. This may best be considered when patient-level data are available. The distribution of an ICER based on sampling error of the numerator and denominator is somewhat complicated, as the 95% confidence interval of a ratio is not easily defined. Nonetheless, there are approaches to understanding the distribution of the ICER. A popular one is to examine the confidence intervals of cost and effectiveness by sampling from the distribution of each, an approach called bootstrap analysis. By sampling both the cost and effectiveness distributions concurrently, multiple estimates of the ICER can be made. These may then be displayed in a plane (Figure 1) in which each point is an estimate of the ICER. In quadrant A, the new therapy is more effective but less costly than the previous standard. In quadrant B, the new therapy dominates the standard, being more effective and less expensive, whereas in quadrant D, the new therapy is dominated by the standard, being less effective and more expensive. The individual ICER estimates can also be used to construct a cost-effectiveness acceptability curve (Figure 2), in which a range of willingness-to-pay thresholds is displayed on the x axis and the y axis indicates the probability that the new therapy has an ICER below this threshold. Thus, the threshold value of the ICER can be varied, and the posterior probability of the ICER meeting that threshold may be noted, a fundamentally Bayesian approach. As the threshold value rises, the probability of the ICER being judged cost-effective increases.

Cost-effectiveness analysis will almost always include a series of assumptions, as it is generally not possible to measure everything necessary for a comprehensive analysis. In addition, even when measurements are available, they may not adequately represent values appropriate for the analysis at hand. Thus, cost-effectiveness analyses generally include sensitivity analyses in which the input variables for assessing both cost and effectiveness are varied. Cost-effectiveness acceptability curves may then be created for each value, or probabilistic Bayesian analyses may be performed, in which all variables affecting cost and effectiveness are varied.

In the approach outlined above, there is no mention of some threshold below which an ICER must be for a new
therapy to be considered cost-effective. The $50,000 per quality-adjusted life-year threshold has been widely used, as it is based on renal dialysis, and in the United States, there is general (political) agreement that there is willingness to pay for renal dialysis. Having a threshold also gives readers of cost-effectiveness studies a benchmark that they can use to judge other studies. However, there is no scientific justification for any one threshold above all others; indeed, the optimal threshold for cost-effectiveness is more a sociopolitical decision than a medical or scientific one.

Diamond and Kaul criticize the $50,000 threshold for judging a new therapy as unrealistic, not generalizable, and poorly defended scientifically. We agree but argue that this is a straw man. In fact, scientists involved in the field do not find much use for the threshold value and pay much more attention to cost-effectiveness acceptability curves. In the absence of a clearly defined threshold, can policy makers use cost-effectiveness analyses? In the Oregon experiment, cost-effectiveness analyses were used to determine which medical therapies would be given.7,8 The cost-effectiveness of therapies were lined up in league tables, and resources were given to therapies until a threshold was reached in which the money gave out. We agree with Diamond and Kaul that such use of cost-effectiveness to set, as opposed to inform, policy is inappropriate. First, this means that all methods used across many clinical areas were comparable and that the comparator was meaningful. It would be a rare person in the field who would make such a claim. It is not at all clear that an expensive therapy with a high ICER given to a child that will offer decades of life can be meaningfully compared with an inexpensive, marginally effective therapy with a low ICER given to an elderly person. Society may choose to offer the former and forgo the latter. Cost-effectiveness analysis can inform this choice but certainly cannot decide the outcome independently of other information. Thus, cost-effectiveness can inform but not decide issues of equity. Using ICERs to decide on access would certainly seem to be inappropriate. Thus, Diamond and Kaul ask, “Should we deny effective but expensive treatment to a child simply because she (or her family) earns less?” Such use of an ICER has nothing to do with the science that lies behind its construction and would indeed seem ethically inappropriate. Diamond and Kaul note the idea of Garber and Phelps9 of tying the ICER threshold to family income. We view this as only food for thought and find it, at best, ethically uncertain. The argument that the cost-effectiveness threshold in poorer areas or countries varies may be correct, although it is arguable. However, even if this argument were agreed upon, it does not follow that cost-effectiveness is not useful. In less wealthy areas where finding value in health care is of greater importance than for the wealthy, cost-effectiveness analysis may be of greater, not lesser, importance. Thus, although cost-effectiveness analysis as developed for medical therapy cannot resolve major societal decisions, it can help poorer counties find value in health, given whatever resources are available for medical care.

There are real problems with cost-effectiveness analysis, which deserve mention. The first is with the quality of data. If a cost-effectiveness analysis is based on 1 or more

Figure 1. Distribution of cost-effectiveness. Each point represents an estimate of the ICER.
randomized clinical results, it will only be as good as the data in the trial. If the trial is biased in some way or not adequately generalizable, the cost-effectiveness analysis will suffer from these same limitations. If a cost-effectiveness analysis is based on a disease simulation model rather than a clinical trial, it will only be as meaningful as the input values. It is also necessary to have an appropriate control group for comparison. Ideally, the control group should represent the current standard of care, assuming that this standard is, itself, reasonably cost-effective. If an inappropriate control group is chosen, the resulting comparison will not lead to efficient resource utilization. Unfortunately, clinical trials of new therapies are often driven by regulatory concerns rather than by addressing important issues of healthcare policy or medical decision making. Finally, the time horizon of a cost-effectiveness analysis may extend beyond the data that are available, requiring modeling of outcome as opposed to direct measurement.

We generally agree with Diamond and Kaul concerning the potential limitations of the $50,000 threshold, but these problems are much less important if the $50,000 figure is used as a guide for understanding value rather than as a willingness-to-pay barrier. We agree that for ICERs, life-years and quality-adjusted life-years should not be used interchangeably, but we find this less of a problem than do Diamond and Kaul. One approach is to present ICERs with both types of denominators. Diamond and Kaul also criticize the Institute of Medicine report that offers several measures of cost-effectiveness. We also do not see this as a problem, as it is best to look at cost-effectiveness issues in several different ways rather than to seek a single number for the ICER that could be compared with the $50,000 threshold. Diamond and Kaul critique the lack of inflation adjustment. This problem is also mitigated by a more nuanced understanding of the $50,000 threshold. Diamond and Kaul note that “perhaps we are taking this fixed $50,000 threshold too literally.” It would seem that they are doing just that.

Some of Diamond and Kaul’s arguments are more philosophical. For instance, they state that “all assessments of value are inherently subjective.” Although we are not entirely sure what they mean by this, it would seem fair to say that although there will be varying degrees of uncertainty in all cost-effectiveness analyses, the ICER is fundamentally objective. Assessment of value or willingness to pay that may follow a cost-effectiveness analysis may include both subjective and objective components. Diamond and Kaul also find circularity in reasoning behind cost-effectiveness analysis. We do not agree. If one were to state that therapy X was cost-effective and therefore it was cost-effective, this would be circular. However, if therapy X has an ICER that is similar to or less than others in a therapeutic area for a similar population, then it would not be circular to argue that it is cost-effective. Diamond and Kaul note that as therapy becomes asymptotic to perfection, new therapy becomes less likely to be cost-effective. We agree and find this a strong argument for cost-effectiveness analysis, because as the benefit decreases toward zero, the ICER climbs toward infinity, which would not meet anybody’s willingness-to-pay threshold. Diamond and Kaul note that “the best way to gauge the value of alternative management strategies is by an up-to-date, real-world, local comparison of the multiple factors entering into interaction between cost and effectiveness.” We could not agree more, and we find this an excellent argument for conducting cost-effectiveness studies.

How then should cost-effectiveness be used? We argue that an ICER is useful. Thus, if a new therapy for heart failure had an ICER of $35,000, payers could consider, in light of other characteristics of the patient population and nature of the therapy, whether that could be considered a good value. Cost-effectiveness should be used to understand how a therapy compares with the previous standard in the field. League tables can be used to help place one study into context. Diamond and Kaul criticize the comprehensive league tables available at a Tufts University Web site (https://research.tufts-nemc.org/cear) for making available >2000 ICERs. It is hard to see why this is a problem, as this would seem to be quite useful information; it does not follow that comparing ICERs makes the $50,000 threshold a rigid
barrier. If the $50,000 threshold is used at all, it should be a general guidepost as to where a specific study falls and certainly not an absolute threshold for medical decision making.

As cost-effectiveness analysis is primarily an approach to evaluating competing therapies, it cannot answer all questions regarding therapy and is by no means the only approach that healthcare economists use. Thus, cost-effectiveness analysis by itself cannot answer the questions raised by Diamond and Kaul, including how many patients will be helped, how much will it cost society, and how we will pay. However, cost-effectiveness analysis may help inform decisions, thus helping society make choices offering good value; by no means does cost-effectiveness analysis suggest that larger societal questions are somehow not important.

Cost-effectiveness analysis is an approach that can be used to extend the understanding of efficacy data, which often come from clinical trials. Appropriately applied, it is more valuable than comparisons of cost alone, sometimes called cost-minimization studies, which implicitly suggest equivalence of efficacy. Cost-effectiveness analysis has many limitations, which are well recognized and clearly enumerated in their article. This fact does not mean that cost-effectiveness analysis is irrelevant or should not be used. It simply means that the “consumers” of this information (ie, guideline committees, policy makers, and health insurers) should be properly educated as to its strengths and limitations. One would not argue that efficacy alone should be the sole criterion by which a therapy should be evaluated; many other issues may limit the value of a therapy, including safety, adverse effects, and the mundane issue as to whether the therapy can be practically and widely applied. Similarly, no one would reasonably argue that cost-effectiveness should be the sole criterion for deciding whether a therapy should be implemented or covered by health insurance. Cost-effectiveness analysis is a method for describing the efficiency by which any therapy leads to improved health. As outlined by Kaul and Diamond, many other factors should be considered in rendering such decisions.

Diamond and Kaul do not answer the question they raise of how we will pay for medical care, for we cannot have all that we would like and thus must seek value. Cost-effectiveness does much more in evaluating value than offer a ratio of cost per outcome gained and the distribution of the numerator and denominator of the ratio. Properly applied, cost-effectiveness analysis can render explicit all of the assumptions underlying the analysis—what are the costs of the therapy, the disease, and its complications; what is the absolute incidence of outcome events and the relative benefit of a therapy; and how do patients and society value these specific outcomes? Whether one chooses to use the cost-effectiveness ratio to represent the inevitable trade-off between costs and effectiveness or to simply report the incremental costs and benefits separately, the fundamental benefits of the exercise are unchanged. An ICER follows this process and is only meaningful when the underlying analytic steps to understand cost and outcome are soundly developed. By forcing physicians, patients, and society to see both sides of the coin, we believe that cost-effectiveness analysis should not be seen as something that is good or bad, but rather as a scientific approach that can help us understand the value of new therapy and thus help inform both medical decision making and public policy. Diamond and Kaul note that the most effective therapy may not be the most cost-effective. True, but suppose a new therapy is 20% more effective than the previous standard at 100 times the cost? Cost-effectiveness can help resolve whether expensive new therapy, while claiming to offer greater effectiveness, also offers good value.

Disclosures

None.

References


Key words: cost-benefit analysis ■ epidemiology ■ cost-effectiveness ■ outcomes