Cost, Effectiveness, and Cost-Effectiveness

George A. Diamond, MD; Sanjay Kaul, MD

Pluck the goose so as to obtain the most feathers with the least hissing.

—Jean-Baptiste Colbert, Minister of Finance to King Louis XIV of France

Incremental or marginal cost-effectiveness ratios are founded on a number of assumptions that weaken their suitability as a way to balance competing economic and clinical priorities. We therefore propose a more relevant and responsible way to gauge the impact of clinical management strategies based on “consumer protection” principles—explicit disclosure of (1) the total magnitude of expected benefit in the target population (in life-years or quality-adjusted life-years), (2) the total monetary cost (in per capita inflation-adjusted dollars), and (3) a formal plan by which the added costs would be paid. Health policy decisions should therefore be based on the inherent trade-offs in the component measures of cost and effectiveness and not on a simple ratio of the two. Cost-effectiveness ratios are thereby rendered superfluous.

Incremental or marginal cost-effectiveness ratios are widely viewed as rational ways to balance competing clinical and economic priorities that arise as a consequence of the inevitable disconnect between an individual’s wants and the society’s willingness to pay for those wants.1–5 The purpose of this essay is to question the practical relevance of these ratios with respect to strategic planning in health care, and offer a suitable alternative. In doing so, we will focus on the more pragmatic issues underlying cost-effectiveness analysis, and gloss over a variety of technical details such as the difference between a privately financed free market and the publicly financed healthcare market, between costs and charges, and between unadjusted and quality-adjusted outcomes.

Our pragmatic perspective is not likely to be welcomed by orthodox experts in cost-effectiveness analysis. We believe, however, that any such criticisms can be blunted by recognizing the important distinctions between regulatory and clinical decisions. As clinicians writing to a clinical audience, we will emphasize the role of cost-effectiveness in clinical decision making throughout much of our discussion, but we will comment on its regulatory role as well.

At its core, cost-effectiveness analysis is no less than a utilitarian moral calculus—one that quantifies the value of any action (even the plucking of a goose) as the ratio of the observable bad qualities, the costs, to the observable good qualities, the benefits. For any action, the lower the ratio of cost to benefit (of hissing to feathers), the higher the value of that action. Although we take no issue with the sophisticated methods used to compute the component values of cost and effectiveness, we will show that the conventional cost-effectiveness ratio suffers from a number of shortcomings as a summary statistic and decision criterion.

We are all familiar with such measures. The price-earnings ratio, for example, values a company in terms of the street price of its stock in proportion to its actual or projected earnings per share. Similarly, the lipoprotein ratio quantifies cardiovascular risk in terms of the ratio of (bad) low-density lipoprotein cholesterol to (good) high-density lipoprotein cholesterol. In the typical healthcare application, the amount of “bad” associated with any action is usually quantified in terms of monetary costs, and the amount of “good” is quantified in terms of the savings in lives, life-years, or quality-adjusted life-years.

The conventional threshold of cost-effectiveness is most often taken to be $50 000 per quality-adjusted life-year (just as the conventional threshold for statistical significance is 0.05). Actions valued below this threshold are commonly considered “cost-effective” and those above are not.6–23 The justification for this singular boundary dates back to a 1980 report on Medicare coverage for treatment of end-stage renal disease, which projected the number of such cases would stabilize at approximately 90 000 patients in 1995 at an inflation-adjusted cost of $4.5 billion.24 This translates to a ratio of $50 000 per life-year (unadjusted for quality). In fact, the numbers have turned out to be a little off. Actual enrollment for 1995 almost tripled (251 214 patients) and the inflation-adjusted cost almost doubled ($8.8 billion)—equivalent to a ratio of only $35 154 per life-year.

In any event, numeric benchmarks such as this are a gross oversimplification of a highly complex process. All assessments of value are inherently subjective, and 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 the interaction between cost and effectiveness. Nevertheless, if we choose to short-circuit this process so as to make simple categorical judgments regarding cost-effectiveness—the canonical goal of cost-effectiveness—
analysis—we must reference these judgments to some defining threshold, even if that threshold is innately naïve.

Thus, a number of liberties have been taken in applying the naïve $50,000 threshold to formal cost-effectiveness analyses on the unlikely assumption that cost and effectiveness are linearly correlated—that if we spend twice as much on something (a meal, a car, a doctor, or a test), we expect it to be twice as good. Medicare’s implied willingness to pay this particular amount for the life-saving treatment of a fatal disease decades ago (motivated largely by political and emotional considerations underlying the Social Security Amendments of 1972, which enacted the end-stage renal disease program) has come to be interpreted, ever since, as the putative valuation for 1 year of a person’s life, and generalized to any and all years, ages, diseases, tests, and treatments. Moreover, life-years have been used interchangeably with quality-adjusted life-years, and studies performed decades apart have used the same fixed threshold without adjustment for inflation.

Although medical economists might well consider these criticisms to be directed at a straw man, saying they really do not believe there is anything special about any particular cost-effectiveness ratio, a casual perusal of their publications leads one to conclude that they certainly act as if they do. Why else does Tufts-New England Medical Center maintain a registry of >2000 cost-effectiveness ratios, all standardized to 2002 dollars? Why else has the $50,000 threshold been used in >250 studies indexed by PubMed—their advocates calling it “economically attractive,” even though it is more than twice the maximum annual Social Security payment of $23,268, and almost 5 times the current annualized minimum wage of $10,712? If the economists do not actually use these thresholds in making their judgments, why do they persist in calculating and publishing them, thereby deceiving clinicians into believing that they indeed rely on them?

Perhaps we are taking this fixed $50,000 threshold too literally. After all, a number of treatments with cost-effectiveness ratios far in excess of this threshold have been accepted into routine clinical practice. Moreover, Garber and Phelps used expected utility theory to derive operative cost-effectiveness ratios in the range of $50,000 or less. Therefore, $50,000 is cost-effective. Renal dialysis costs $50,000. Renal dialysis is cost-effective. Therefore, $50,000 is cost-effective.

This is, in fact, a valid argument. The truth of the conclusion follows from the antecedent premises. But the premise that renal dialysis is cost-effective is itself assumed, and the entire argument therefore begs the question—assuming just what is to be proven.

It would seem self-evident that any cost-effective strategy should be transportable—what’s good for Californians ought to be good for Texans—but the conventional threshold violates this principle. According to Table 1, for example, the $50,000 threshold is larger than the annualized per capita income in the most affluent state in the country and is twice that in the least affluent state. As a result, a strategy deemed to be cost-effective on a national basis will not be cost-effective in

<table>
<thead>
<tr>
<th>National Ranking</th>
<th>State</th>
<th>Per Capita Income, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Connecticut</td>
<td>49,852</td>
</tr>
<tr>
<td>25</td>
<td>Texas</td>
<td>34,257</td>
</tr>
<tr>
<td>50</td>
<td>Mississippi</td>
<td>26,535</td>
</tr>
<tr>
<td></td>
<td>US average</td>
<td>36,276</td>
</tr>
</tbody>
</table>

Table 1. Per Capita Income for Selected States

Table 2. Per Capita Income (Purchasing Power Parity) for Selected Countries

<table>
<thead>
<tr>
<th>World Ranking</th>
<th>Country</th>
<th>Per Capita Income, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/181</td>
<td>United States</td>
<td>45 845</td>
</tr>
<tr>
<td>15/181</td>
<td>Canada</td>
<td>38 435</td>
</tr>
<tr>
<td>61/181</td>
<td>Mexico</td>
<td>12 775</td>
</tr>
<tr>
<td></td>
<td>World Median</td>
<td>6964</td>
</tr>
</tbody>
</table>


the majority of individual states. Applying Garber and Phelps’ doubling rule to the data in Table 1, for example, one would conclude that treatments possessing cost-effectiveness ratios between $50 000 and $100 000 should be offered to residents of Connecticut but not Mississippi (!).

What about similar geographic variations in pricing? Because the government does not compute the Consumer Price Index on a state-by-state basis, it is reasonable to presume that geographic price variability parallels geographic income variability. As a result, Garber and Phelps justifiably adjust cost-effectiveness ratios in terms of per capita income alone. It is unlikely that the relatively small geographic variation in medical pricing independent of that reflected in per capita income variations or the modest leveling effects of national reimbursement programs such as Medicare can help resolve this problem.

The conventional $50 000 threshold has even less relevance to other countries, most of which are not as well off as ours. As shown in Table 2, again using per capita income (adjusted for the economists’ standard of purchasing power parity) as the reference, a strategy that is cost-effective in the United States is not likely to remain so beyond our borders. The World Health Organization recognizes this problem, but its proposed solution—defining the nation-specific threshold as 3 times its per capita income without regard for purchasing power parity—is nothing of the sort. How many impoverished individuals in the Third World would be willing to spend (or have their government spend) 3 times their meager annual income on a single medical procedure instead of providing them with clean water and effective sanitation? In fact, we would have to cut the threshold by 80% just to maintain parity with the rest of the industrialized world where the basic public health infrastructure is already in place.

Even if we addressed all these details of social calibration, cost-effectiveness would nevertheless remain inadequate to the task at hand. Who really wants to maximize cost-effectiveness as the priority in making medical decisions? Who would choose the most cost-effective treatment for themselves or their families instead of the most effective treatment? Here’s a concrete example. Systemic medical therapy (in the form of aspirin, statins, nitrates, β-blockers, and calcium-channel blockers) is arguably more cost-effective than local revascularization for the management of most forms of chronic stable angina given the latter’s higher cost and lower effectiveness (owing to the attendant risks of death, myocardial infarction, restenosis, and atherosclerotic progression). Nevertheless, revascularization is widely viewed as the preferred treatment option. Although healthcare payers are cognizant of these cost-effectiveness distinctions, they readily reimburse physicians for the more costly revascularization procedure and increase their premiums accordingly. Patients and employers willingly pay these higher premiums on the mistaken assumption that the availability of high-tech care is equivalent to the provision of high-quality care. In fact, Medicare does not consider cost-effectiveness at all in deciding reimbursement issues, and we would very likely object loudly if it did.

In excluding cost-effectiveness as a criterion for regulatory decision making, Medicare is acting in full compliance with operative federal law, and specifically, with Executive Order 12866 enacted by President Clinton on September 30, 1993:

> In deciding whether and how to regulate [federal] agencies should assess all costs and benefits of available regulatory alternatives, including the alternative of not regulating. Costs and benefits shall be understood to include both quantifiable measures (to the fullest extent that these can be usefully estimated) and qualitative measures of costs and benefits that are difficult to quantify, but nevertheless essential to consider. Further, in choosing among alternative regulatory approaches, agencies should select those approaches that maximize net benefits...

Although the term “net benefit” has a formal meaning to economists (total benefit minus total cost), it is used here in a more casual way. The only reference to cost-effectiveness in this document refers not to the decision to regulate but to the design of the regulation once such a decision is made:

> When an agency determines that a regulation is the best available method of achieving the regulatory objective, it shall design its regulations in the most cost-effective manner to achieve the regulatory objective [so] that the benefits of the intended regulation justify its costs.

Nevertheless, some argue that considerations of cost-effectiveness be incorporated into policy decisions regarding health insurance coverage. In 2006, the Institute of Medicine issued a comprehensive report that walks a tightrope on the matter. Thus, while acknowledging the priority of overall benefits and costs, consistent with the executive order, it recommends also that “regulatory analyses should report 4 measures of cost-effectiveness”:

- gross cost per life saved,
- gross cost per year of life saved,
- gross cost per quality-adjusted year of life saved, and
- net cost per quality-adjusted year of life saved.

Such indecisiveness can lead to contradictory conclusions (see Alternative Cost-Effectiveness Criteria). Therefore, rather than advising regulators, economists, and insurers to incorporate specific cost-effectiveness criteria into their policy decisions, we propose empowering the end-users—physicians and patients. Accordingly, in the spirit of the executive order, we posit 3 questions based on accepted principles of consumer protection that better address the justification of any proposed healthcare strategy, questions
that any socially responsible healthcare advocate should be expected to answer:

1. How many people will it help? What is the additional number of (quality-adjusted) life-years saved, or events prevented in the total population under consideration? The larger that number, the better the strategy.

2. How much will it cost? The lower the cost, the better the strategy. As above, the additional cost should be referenced to the total population under consideration. In purchasing a new car, we want to know the true cost of our purchase, not the cost per month or the cost per mile that the salesperson tries to promote to us in a veiled attempt to hide the true cost from us. Anyone wishing to compute a conventional cost-effectiveness ratio (relative to some suitable alternative) can do so with these 2 numbers, but it is really superfluous. The raw numbers themselves communicate all we need to know.

3. Perhaps most important of all, how do we plan to pay for it given competing needs and existing budgetary constraints? We cannot afford blanket coverage for all the “cost-effective” care we might like, and trimming the wasteful “cost-ineffective” care will never make up the difference.

Conventional cost-effectiveness ratios can answer none of these questions. In contrast, Table 3 summarizes a simple “back of the envelope” calculation for application of our consumer-protection approach to cardiovascular prevention. In this example, we compare an unconditional treatment strategy (treat everyone) against a conditional test-treatment strategy (screen everyone with a test, and treat only those with an abnormal response). Let us assume that the “treatment” is a preventive drug such as a statin, at a cost of about $2 per day ($720 per year), and the “test” identifies an “at-risk” population experiencing 4% of the events (400 000 events, which is an event rate of 1%).

If we treat every one of these adults, we can expect to reduce these events by 30% based on available randomized clinical trials. We will thereby prevent 150 000 events at a total cost of $36 billion (90% of 50% of 50 million patients). The alternative strategy would have us test all 50 million individuals at $100 per patient. On the basis of Pareto’s 80/20 rule, we can expect our test to identify 20% of the population (10 million patients) who will experience 80% of the events (400 000 events), which is an event rate of 4%. If we now treat only this higher risk population, we will prevent 120 000 events at a total cost of $12.2 billion ($5 billion for testing plus $7.2 billion for treatment). The unconditional treatment strategy therefore costs an additional $23.8 billion but prevents 30 000 more events. Assuming each event represents a loss of 12.9 life-years, the marginal cost-effectiveness of the unconditional treatment strategy is $61.500 per life-year ($23.8 billion divided by 30 000 events divided by 12.9 life-years per event). The more effective strategy (unconditional treatment) is therefore not cost-effective using the conventional $50 000 threshold.

Of course, different parameters would materially change this conclusion. If the cost of treatment could be reduced to $610, for example, the cost-effectiveness ratio for the unconditional strategy would fall below $50 000. Alternatively, if the tests were interpreted using a more discriminatory diagnostic criterion (eg, capable of identifying 10% of the population experiencing 70% of the events), the cost-effectiveness ratio for the unconditional strategy would again fall below $50 000.

But is that really the point? Whatever the costs (and resultant cost-effectiveness ratios), in the end, the unconditional treatment strategy always prevents more events. Testing might save money, but it cannot save lives; only treatment saves lives. We will leave it to the advocates of the alternative testing strategy to justify this trade-off between cost and benefit, and tell those who might otherwise have been saved that it just was not worth the added expense, and that it was, dare we say, more cost-effective to let them go.

In the final analysis, the key issues involved in cost-effectiveness decisions are matters of politics, not of science. As Garber and Phelps44 phrased it in an earlier version of their seminal analysis of the economic foundations of cost-effectiveness analysis:

The optimal [cost-effectiveness] ratio...is quite sensitive to income and attitudes toward risk. If a single...ratio is applied to all interventions and to all individuals in a group, for some of them the marginal benefit will fall much lower than the marginal cost, and for others, just the opposite….[Cost-effectiveness] analysis applied at the population level may give the most efficient egalitarian distribution of health resources, but it is not likely to be Pareto optimal.

In other words, what’s good for the group is not necessarily good for individual members of the group. Whether for health care or national defense, a proper accounting of marginal costs and benefits—not their ratio—is the more sensible starting point for strategic decisions from a societal perspective. Only patient-specific costs and patient-specific benefits will have relevance to clinical decisions regarding individual patients. As Monsieur Colbert might have known, we are best advised to focus more on long-term net worth (the absolute difference between assets and liabilities) than on short-term rate of return (the ratio of profit to loss).

### Alternative Cost-Effectiveness Criteria

The Panel on Cost-Effectiveness in Health and Medicine was convened a decade ago to establish standards and resolve controversies for the performance of cost-effectiveness anal-

<table>
<thead>
<tr>
<th>Metric</th>
<th>Treat</th>
<th>Test→Treat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target population</td>
<td>50 million</td>
<td>10 million</td>
</tr>
<tr>
<td>Expected events</td>
<td>500 000</td>
<td>400 000</td>
</tr>
<tr>
<td>Annual cost</td>
<td>$36 billion</td>
<td>$12.2 billion</td>
</tr>
<tr>
<td>Prevented events</td>
<td>150 000</td>
<td>120 000</td>
</tr>
</tbody>
</table>
yses across a spectrum of medical disciplines.\textsuperscript{1,3} Their 1996 report proposed a number of technical methodological guidelines, but specifically avoided recommending a particular cost-effectiveness ratio as a formal decision-making threshold. Although the report had some impact on the subsequent quality of cost-effectiveness analyses, the majority of recommendations have been ignored.\textsuperscript{46}

In a more recent report, the Institute of Medicine recommends 4 different definitions of cost-effectiveness, again sidestepping the recommendation of a particular decision threshold.\textsuperscript{36} Multiple definitions of effectiveness (and to a lesser extent, cost) can be justified by the fact that the anticipated benefit associated with some forms of treatment is naturally referenced to quality of life (eg, cataract surgery), whereas that of others is more naturally referenced to survival (eg, cancer surgery), and still others might be better referenced to quality-adjusted survival (eg, cancer chemotherapy).

According to the first definition, 2 treatments (A and B) are equivalent when the cost per life saved for each is equal. For purposes of illustration, we can represent this condition by the following equation:

$$\frac{c_A}{m_0 - m_A} = \frac{c_B}{m_0 - m_B}$$

where $m_0$ is the annual mortality rate associated with non-treatment, $c_A$ and $m_A$ are the annual cost and mortality rate, respectively, associated with treatment A, and $c_B$ and $m_B$ are the annual cost and mortality rate, respectively, associated with treatment B. Thus, if $m_0 = 0.12$, $c_A = $10 000, $m_A = 0.10$, and $m_B = 0.07$, treatment B will be more cost-effective than treatment A if it costs less than $25 000 ($500 000 per life saved).

According to the second definition, 2 treatments are equivalent when the cost per year of life saved for each is equal:

$$\frac{c_A}{m_A} = \frac{c_B}{m_B}$$

where the reciprocal of each annual mortality approximates the attendant life expectancy,\textsuperscript{47,48} and the “years of life saved” are thereby estimated as the marginal difference in life expectancy for each treatment. Using the same values as before, treatment B will then be more cost-effective than treatment A if it costs less than $35 714 ($6000 per year of life saved).

According to the third definition, 2 treatments are equivalent when the cost per quality-adjusted year of life saved for each is equal:

$$\frac{c_A - s_A}{m_A + q \mu_A} = \frac{c_B - s_B}{m_B + q \mu_B}$$

where $s_A$ and $s_B$ are the indirect cost savings associated with each treatment (as a consequence, for example, of return to work). If $s_A = $5000 and $s_B = $3000 while all other values remain the same, treatment B will then be more cost-effective than treatment A if it costs less than $9417 ($1558 net per quality-adjusted year of life saved). Thus, because one’s choice of the operative definition is arbitrary, so too is one’s judgment of cost-effectiveness.

Disclosures

None.

References


35. Federal Highway Administration. Evaluation and determination for Federal-aid transportation projects that have a net benefit to a Section 4(f) property. *Federal Register*. 2005;70:20618–20630.


**Key Words:** cost-effectiveness ■ health policy ■ medical economics