Recommendations of the Panel on Cost-Effectiveness in Health and Medicine

Milton C. Weinstein, PhD; Joanna E. Siegel, ScD; Marthe R. Gold, MD, MPH; Mark S. Kamlet, PhD; Louise B. Russell, PhD; for the Panel on Cost-Effectiveness in Health and Medicine

Objective.—To develop consensus-based recommendations for the conduct of cost-effectiveness analysis (CEA). This article, the second in a 3-part series, describes the basis for recommendations constituting the reference case analysis, the set of practices developed to guide CEAs that inform societal resource allocation decisions, and the content of these recommendations.

Participants.—The Panel on Cost-Effectiveness in Health and Medicine, a nonfederal panel with expertise in CEA, clinical medicine, ethics, and health outcomes measurement, was convened by the US Public Health Service (PHS).

Evidence.—The panel reviewed the theoretical foundations of CEA, current practices, and alternative methods used in analyses. Recommendations were developed on the basis of theory where possible, but tempered by ethical and pragmatic considerations, as well as the needs of users.

Consensus Process.—The panel developed recommendations through 2 1/2 years of discussions. Comments on preliminary drafts prepared by panel working groups were solicited from federal government methodologists, health agency officials, and academic methodologists.

Conclusions.—The panel's methodological recommendations address (1) components belonging in the numerator and denominator of a cost-effectiveness (C/E) ratio; (2) measuring resource use in the numerator of a C/E ratio; (3) valuing health consequences in the denominator of a C/E ratio; (4) estimating effectiveness of interventions; (5) incorporating time preference and discounting; and (6) handling uncertainty. Recommendations are subject to the "rule of reason," balancing the burden engendered by a practice with its importance to a study. If researchers follow a standard set of methods in CEA, the quality and comparability of studies, and their ultimate utility, can be much improved.

JAMA. 1996;276:1253-1258

COST-EFFECTIVENESS analysis (CEA) has emerged as a basic tool in the evaluation of health care practices. Despite widespread application, there remain disparities in the methods that investigators employ. Some of these disparities can be traced to a misunderstanding of the principles of CEA, while others reflect divergent views on key methodological choices. For example, an investigator who fails to account for important negative side effects of a therapy in estimating effectiveness is making a clear error, while investigators who include or exclude the financial costs of lost productivity that accompany illness are reflecting different views of how productivity should be accounted for in a CEA.

The divergence of methods used to conduct CEA interferes with the ability of decision makers charged with resource allocation to make appropriate comparisons of cost-effectiveness (C/E) ratios across programs. As described in the first article of this series, this concern about lack of standardization has led the Panel on Cost-Effectiveness in Health and Medicine to develop a set of recommendations for the practice of CEA that can serve as a point of reference for investigators who seek comparability with other analyses in the literature. The panel refers to this set of methodological practices as the reference case.

The reference case will not address all types of questions regarding the cost-effectiveness of interventions. In some cases, depending on the goals of the analysis, the author may prefer to highlight an analysis based on a slightly different set of principles, or one based on quite different assumptions. In the interest of comparability, however, we urge that the reference case set of assumptions and practices be included in every CEA that is designed to permit broad comparisons across interventions or that might be used for this purpose.

The recommendations outlined here, together with others that provide more detailed methodological guidance, are expanded in the full report of the panel. While this article focuses on the reference case recommendations, we also describe a few recommendations that are intended to improve the conduct of analyses but that are not explicitly incorporated within the reference case.

RATIONALEs FOR RECOMMENDATIONS

Reference case analyses are intended to inform resource allocation decisions and, as described in the first article of this series, are conducted from the societal perspective for this reason. Specific rec-
ommendations for conducting CEs from the societal perspective were based on a number of considerations, including economic and decision theory; consistency in the accounting of costs and consequences; ethical concerns; pragmatic concerns; and needs of users of analyses. In some instances, where neither theory nor these other considerations led to a clear choice, the panel recommended a conventional practice for the sake of consistency across studies.

Where possible, recommendations were based on theoretical considerations in order to provide a defensible, logical, and consistent framework for methodological choices. At one level, CEs can be based solely on the mathematical theory of “optimization.” In that formulation, the decision maker is free to select any objective to maximize (eg, life-years or quality-adjusted life-years [QALYs]) and to specify the particular resource constraint under which allocations must be made (eg, national health care costs). However, this framework alone provides no guidance on key issues that arise from a societal perspective, such as which costs to include in the analysis, how to measure costs, and whose values to incorporate into the definition of health consequences. Therefore, drawing from recent literature suggesting a link between CEs and welfare economic theory, the panel relied on economic theory for many of the recommendations for the reference case. In addition to economic theory, principles of decision theory were invoked to define the basis for individual preferences.

Some recommendations were dictated by the need to maintain a logical accounting of costs and health effects. For example, theoretical considerations, along with a basic presumption about the definition of the C/E ratio, led to a clear decision concerning the placement of all health effects in the denominator of the C/E ratio. Then, the “accounting” principle that no cost or effect should be counted twice disallowed the inclusion of health effects—even in monetary form—in the numerator of the C/E ratio.

The implications of welfare economics were often modified in the interest of producing recommendations that were both pragmatic and ethically acceptable. The practical need to obtain data on health outcomes, utilization of services and unit costs, and weights for health-related quality-of-life (HRQL) states led to such compromises. For example, while medical prices are not an exact reflection of the true value of resources, pragmatism suggests that prices be used to approximate costs except where distortions are likely to be significant and important to the analysis.

Ethical considerations sometimes tempered recommendations based on economic theory or were decisive in choices among alternatives. Most fundamentally, the decision to use QALYs as the effectiveness measure reflects the ethical stance that QALYs accruing to different people or at different stages in life should be valued equally, even though welfare economics implies that health benefits should be weighted by willingness to pay.

Where theory did not offer a clear choice, the panel based some recommendations simply on the need for a clear conversion to which analysts would adhere in the reference case. In some cases, the recommendation was somewhat arbitrary. For example, the choice of a standard time discount rate, while guided by theory and data, is fixed by the need for a standard practice.

Finally, needs of the potential users of CEs influenced several recommendations, playing a particularly great role in the panel’s recommendations for the reporting of CEs as described in the third article of this series. They also entered into the recommendations regarding the evidence of effectiveness used in analyses and the treatment of uncertainty. For example, sensitivity analyses, which explore the implications of alternative assumptions and data, are often recommended so that decision makers can gain confidence in the conclusions of an analysis or identify areas for further investigation.

While some of the reference case recommendations address common errors in the practice of CEs, many more represent the panel’s view of the best among several defensible choices. The task of developing CEs standards is analogous to the formulation of the consumer price index (CPI), which is used to adjust for inflation based on the prices of a typical market basket of goods and services. The choice of what items go into that market basket, and how they are weighted, reflects judgments made by the Bureau of Labor Statistics and its advisors. Legitimate opposing views exist. However, there is an implied consensus that the CPI will be used so that industry, government, and consumers can have a shared understanding of the inflation rate. Indeed, flows of resources, such as the level of Social Security payments, depend in part on the CPI. While this country does not base policy directly on C/E ratios as it does with the CPI, it is important for many decision makers to be able to rely on a dependable yardstick for measuring cost-effectiveness of health services.

RECOMMENDATIONS

The panel’s recommendations fall largely into 8 categories: (1) the nature and limits of CEA and of the reference case; (2) components belonging in the numerator and the denominator of a C/E ratio; (3) measuring terms in the numerator of a C/E ratio (costs); (4) valuing the health consequences in the denominator of a C/E ratio; (5) estimating effectiveness of interventions; (6) time preference and discounting; (7) handling uncertainty in CEA; and (8) reporting guidelines. The first group of recommendations, regarding the nature and limits of CEA, has been described in the first article of this series. The last group, regarding reporting guidelines, is the subject of the third article. Additional recommendations regarding research to develop improved data for CEA and improved methods are described in the full report of the panel.

This article summarizes the remaining 6 categories of recommendations.

Components Belonging in the Numerator and the Denominator of a C/E Ratio

Cost-effectiveness analysis rests on the proposition that a decision maker wishes to select programs so as to maximize some desired objective subject to a resource constraint. In practice, CEA in health care has been based on the premise that health benefits are the objective that societal decision makers wish to maximize, subject to a constraint on health care resources. This formulation leads directly to the construction of a C/E ratio in which the net expenditure of health care resources (a monetary measure) goes in the numerator and the net improvement in health (a nonmonetary measure) goes in the denominator.

Unfortunately, however, this definition is incomplete. It leaves open to question whether certain costs and consequences should be thought of as health care costs or savings (numerator), or health decrements or improvements (denominator), and it completely ignores nonhealth costs and effects, such as those associated with economic productivity, the environment, or education. Therefore, if analyses are to handle such issues consistently, the choice between numerator (resource impact) and denominator (HRQL impact) must follow an established convention. In any case, the societal perspective dictates that all important impacts on human health and on resources must be included somewhere, either in the numerator or the denominator. With this principle in mind, the panel reached the following recommendations regarding the distinction between costs and health consequences.

By convention, the denominator of a C/E ratio is reserved for the improvement in health associated with an intervention. Thus, effects of an intervention on length of life and on morbidity, including the full value of HRQL to patients, should be incorporated in the denomina-
consideration, but are made for consistency across reference case analyses.

### Measuring Terms in the Numerator of a C/E Ratio (Costs)

A change in the use of a resource caused by a health intervention should be valued at its opportunity cost, which is the value the resource could have produced if it were spent in its best available alternative use. In economics, this principle is the basis for valuing resource flows in society.

Several implications arise from the opportunity cost principle. First, it is the difference in resource use between an intervention and the intervention to which it is being compared that should be included in the numerator of the C/E ratio. That is, costs should reflect the marginal or incremental resources consumed or saved, rather than total resources. Fixed costs—costs unaffected by the level of implementation of an intervention—should generally be excluded from consideration. However, resource costs should be measured from a long-term perspective, which implies that many costs that may be fixed in the short run (such as most of what is usually considered overhead in the financial accounts of hospitals and other health care providers) are in fact variable in the long run and should be included in CEAs.

Direct measurement of opportunity costs is difficult and often impossible. To the extent that market prices of health-care inputs reflect opportunity costs, they provide an appropriate means for valuing changes in resources. According to economic theory, prices in competitive markets reflect opportunity costs of resources. However, when prices do not adequately reflect opportunity costs because of market distortions, they should be adjusted appropriately. Examples of adjustments commonly used in CEAs include the use of ratios of cost to charge (RCCs) to adjust hospital prices, the use of management accounting systems to estimate costs, and the use of third-party payments to providers in lieu of fees to reflect provider opportunity costs. When substantial bias is present in prices and adjustment is not feasible, the panel recommends that more suitable proxies for opportunity costs be considered, including the possibility of conducting "microcosting" studies within provider organizations. (Such studies collect information on the range of inputs to a service, such as the nursing care, supplies, and ancillary services constituting a day of hospital care.)

Costs should be measured in constant dollars, that is, in dollars of a fixed year. When the original data are for different years, the effect of price inflation must be removed, either by inflating the data from an earlier year to the chosen year or by deflating the data from a later year. Depending on whether the resources being valued are more representative of goods and services in the economy at large or in the medical care sector, either the CPI or its medical care component(s) is suitable for inflation adjustment in CEAs.

Transfer payments (such as cash transfers from taxpayers to welfare recipients) associated with a health intervention redistribute resources from one individual to another. While administrative costs associated with such transfers could be included in the numerator of a C/E ratio, the transfers themselves should not be, since, by definition, their impact on the transferor and the recipient cancel out from the societal perspective.

Time costs for individuals in the labor force should generally be valued by the wage rate as an acceptable measure of opportunity cost of time. The reference case recommendation is to use wages corresponding to the age and gender composition of the target population. However, group-specific wages may influence the conclusions of an analysis in ethically problematic ways. For example, a policymaker might object to having the wage differential between men and women reflected in the results of a CEA. In these instances, sensitivity analyses should be conducted to explore the specific nature of this influence. Because of such ethical concerns, and because of practical problems in obtaining data on wages by characteristics other than age and gender, the panel does not recommend using wages specific to target groups defined by race, ethnicity, or other characteristics.

Wage rates generally do not adequately reflect the value of time for persons engaged primarily in leisure—such as retired persons—or in activities for which they are not compensated—such as household activities. Average age- and gender-specific wages among persons in the labor force may be applied to approximate the opportunity cost of time for persons of similar age and gender not in the labor force.

Should health care costs that result solely from the fact that a successfully treated patient lives longer be attributed to the health intervention? Which future costs should be included in a CEA? For example, a cost-effective analysis of antihypertensive therapy found that excluding noncardiovascular disease costs in future years would reduce the C/E ratio by 5% to 20%, with the greatest impacts on the ratios for younger population groups. To clarify the issues, we define 5 categories of induced costs that may or may not be germane in a CEA. These are (1) costs for intervention-related diseases incurred in years of life that would have been lived anyway; (2) costs for unrelated diseases...
that are incurred in years of life that would have been lived anyway; (3) health care costs for related diseases that ensue in years of life added (or subtracted) as a result of the intervention, (4) health care costs for unrelated diseases that occur in years of life added (or subtracted) by the intervention, and (5) nonhealth care costs typified by commodities such as food and shelter that occur in years of life added (or subtracted) by the intervention.

The handling of some of these categories of costs is uncontentious. Costs of related diseases in the original life span clearly must be included in the analysis. For example, costs and savings associated with treatment of strokes and myocardial infarctions must be included in analyses of hypertension programs. Costs of treating adverse effects of treatment must be included as well. On the other hand, because unrelated health and nonhealth care costs occurring throughout the expected years lived without the intervention would cancel from the incremental cost calculation in the numerator of the C/E ratio, these may be omitted from the analysis. It may actually be preferable to omit these unrelated costs because their measurement may add to error in the estimation of costs with and without the intervention.

Costs for intervention-related diseases that occur in added years of life are typically included in CEA. As for example, if a fatal myocardial infarction is delayed 5 years by a coronary bypass operation or a cholesterol-lowering regimen, the costs of treating coronary events ensuing throughout the 5 years should be, and usually are, included. Similarly, costs of an ongoing therapy throughout added years of life, such as lifelong antihypertensive treatment and its medication side effects, are always included.

Costs of diseases unrelated to the intervention and ensuing as a result of added years of life have been the source of more debate. Difficulties with the choice to include or exclude them are illustrated by the example of a cholesterol-lowering intervention. The analyst might decide to exclude all unrelated costs occurring in years of life gained because of the program. In this case, costs of illnesses such as arthritis and Alzheimer disease would be excluded. However, age-specific "background" costs of coronary heart disease—that is, the level of disease that would occur among people who are not candidates for the intervention—are also "unrelated" to the intervention and should also be excluded. To neglect to do so would provide an uneven playing field for comparisons of interventions affecting different diseases: life-prolonging heart disease interventions would be commensurate with all future costs of heart disease even though they only target an excess risk, while suicide prevention programs would not. To avoid the practical and conceptual problems in disengaging the "related" and "unrelated" elements of costs for "related" diseases, it would be preferable to include all these costs. However, this choice would impose a burden on the analyst frequently not warranted by the importance of future costs.

Because there are unresolved theoretical and empirical questions and because health care costs in added years of life are typically small compared with the other costs in an analysis, the panel concluded that the reference case may either include or exclude these costs. Whenever the investigator has reason to believe that inclusion or exclusion of future health care costs may make a significant difference to the analysis, a sensitivity analysis should be performed to assess the effect on the C/E ratio.

We now consider nonhealth care costs in added life-years. Although there is no precedent in CEA for including these costs, one could reasonably argue that if health care costs in added years can be included, future expenditures on food, clothing, and shelter should also be included. The theoretical answer is that the net economic burden of survivors on the rest of society (consumption minus productivity) should indeed be included as a cost. However, if these nonhealth care costs are truly "unrelated," then their consistent inclusion or exclusion would only add or subtract a constant from the C/E ratio. Whether nonhealth care costs are in fact "unrelated," or at least approximately so, is an unresolved empirical question. Nonetheless, on the assumption that these costs can reasonably be considered to be unrelated and to avoid placing an unnecessary burden on the analyst, the panel does not recommend including future nonhealth care costs in reference case analyses.

Valuing the Health Consequences in the Denominator of a C/E Ratio

As discussed in the first article of this series, a reference case analysis should measure health effectiveness in terms of QALYs. These QALYs incorporate changes in survival and changes in HRQL by weighting years of life to reflect the value of the HRQL during each year.

In order to be consistent with the QALY construct, the quality weights must be measured by or transformed into the interval scale on which optimal health has a value of 1 and death has a value of 0.

The weights used in QALYs should be based on a health-state classification system that reflects health-related domains (attributes) that are important for the particular analysis. In order to qualify as a reference case analysis, the CEA should use a generic health-state classification, that is, a classification that applies broadly across diseases and conditions. Disease-specific health-state classifications are appropriate for a reference case analysis provided that they are designed to be mapped onto or embedded within a generic system. Some examples of commonly used health-state classification systems that may be suitable for CEA include the Health Utilities Index, the EuroQol, the Quality of Well-Being Scale, and the Years of Healthy Life measure. The Health Utilities Index, for example, consists of 8 domains (vision, hearing, speech, dexterity, mobility, cognition, emotion, and pain), each of which is classified into either 5 or 6 levels. Each combination of levels is assigned a weight, using a formula based on multiattribute utility theory and a community preference survey.

The weights used in QALYs should be based on preferences for health states. In a reference case analysis, these weights should be based on community preferences, rather than those of patients, providers, or investigators. The rationale for community preferences has been described in the first article of this series. Health status scales that are not preference weighted, such as the Medical Outcomes Study Short-Form Health Survey (SF-36), are not suitable for CEA in their present form. Use of patient preferences to value health states is acceptable in a reference case analysis only when adequate information is unavailable regarding community preferences.

The weights assigned to health states should be interval scaled; that is, the method of measurement should be one in which the ratio of differences between values is meaningful. (By analogy, the Fahrenheit, Celsius, and Kelvin scales are equivalent and appropriate measures of temperature, because the intervals between degrees reflect meaningful differences in temperature.) According to decision theory, preference weights obtained from standard gamble questions and, under certain conditions, time-trade-off questions satisfy the interval property. At the same time, psychometric research suggests that rating scales can produce interval data. These claims are mutually inconsistent, since there is apparently not a linear relationship between rating scales and standard gambles or time trade-offs.

It remains an open question whether standard gambles, time-trade-offs, rating scales, or other measures such as person trade-offs produce the closest approximation to the type of interval-scaled weights needed for QALYs. For example, some research suggests that respondents may introduce distortions
into responses to utility questions such as the standard gamble, compromising their theoretical desirability. Therefore, the panel does not recommend one source of weights over the others; for purposes of the reference case, preference weights can come from measurement systems that rely on any of these techniques. The discrepancies associated with different measurement strategies pose a problem for standardization that will be important to address in future research.

To date, most CEAs using QALYs have assumed that a year of life gained by an intervention is valued at 1.0 QALY. In fact, people are rarely in the state of optimal health that a full QALY implies; the use of a value of 1.0 for years that life is prolonged will therefore overstate an intervention’s effectiveness and underestimate the true C/E ratio. The panel recommends that, when calculating QALYs gained from a life-extending intervention, estimates of age- and sex-specific HRQL should be applied to the years of extended life—even if the intervention itself has no effect on HRQL. Similarly, when estimating QALYs gained by ameliorating disease symptoms, a return to average rather than optimal HRQL should be credited. It should be noted that this use of average quality of life in reference case analyses means that studies using ratios of cost per year (unadjusted) of life saved will not be comparable to reference case results.

Sociodemographic characteristics, such as age, sex, or race, are associated with HRQL. When the QALYs produced by an intervention vary as a result of these sociodemographic differences, reference case results are affected in ways that may be ethically problematic. For example, an intervention that extends the lives of 80-year-olds may appear less cost-effective than an equally effective intervention applied to 20-year-olds, not only because fewer years are gained, but also because the average quality of those years is less. In these instances, sensitivity analyses should be conducted to indicate explicitly how results are affected.

### Estimating Effectiveness of Interventions

The quality and validity of a CEA depend crucially on the quality of the underlying data that describe the effectiveness of interventions and the course of illness without intervention. Data may be obtained from primary data collection efforts specifically intended to inform the CEA or from secondary data sources. The appropriateness of various sources of data will depend on the purpose of a CEA. The consequences of misestimation of cost-effectiveness may be regarded as more serious by some decision makers than by others. For example, the Food and Drug Administration might desire a greater level of certainty in distinguishing the cost-effectiveness of very similar drugs for the purpose of reviewing marketing claims than a formulary manager might demand in adopting a new drug.

For the purpose of a reference case analysis, acceptable data for estimation of effectiveness may come from a variety of sources: randomized controlled trials, observational studies, uncontrolled experiments, or descriptive series. The analyst should select outcome probabilities from the best-designed (and least-biased) sources that are relevant to the question and population under study. There are often trade-offs between the internal validity of data (optimized in randomized trials) and their external validity in actual practice. Meta-analysis and other synthesis methods can be used when single study has sufficient power to detect effects or when studies produce conflicting results. Expert judgment should be used only as placeholders where no adequate empirical data exist, or when the parameter of interest plays only a minor role in the analysis.

Modeling is a valid and necessary scientific procedure for estimating effectiveness for CEA. Typically, data from randomized trials are combined with observational data and public health statistics in models that are used to estimate changes in life expectancy and quality-adjusted life expectancy. Models may incorporate features such as logistic regression to estimate incidence of disease or death contingent on risk factors; Bayesian analysis to estimate posttest probabilities of disease from data on sensitivity, specificity, and prevalence; and life-table analysis to estimate life expectancy from survival curves. Models include population and cohort models, deterministic and probabilistic models, decision analysis and state-transition models. Because of limited time horizons and selected study populations in clinical studies, failure to use models to extrapolate from primary data can lead to greater error than the models themselves would introduce. Models should be used as complements to, not substitutes for, direct primary or secondary empirical evaluation of effectiveness. Readers are referred to the full report of the panel for more discussion of the roles of particular types of data and models in CEA.3

### Time Preference and Discounting

Interest rates reflect people’s preference for having money and material goods sooner rather than later. Similarly, people value health outcomes that occur in different time periods differently. In CEA, time preference for resources is reflected by discounting future costs to present value. Discounting the value of future expenditures requires that health effects experienced in the future also be discounted at the same rate. This conclusion is based on the observation that people have opportunities to exchange money for health, and vice versa, throughout their lives. Failure to discount health effects will lead to inconsistent choices over time; for example, it will appear that delaying investments will always result in a program’s becoming more cost-effective. For this reason and based on other evidence and considerations outlined in its full report,2 the panel recommends that costs and health outcomes occurring during different time periods should be discounted to their present value and that they should be discounted at the same rate.

Although a wide variety of discount rates are used in the literature and can be defended, a convention is needed to achieve consistency across analyses. Theoretical considerations suggest that the real (inflation-adjusted) discount rate should be based on time preference, the difference in value people assign to events occurring in the present vs the future. Further, economic theory suggests that time preference is reflected in the rate of return on riskless, long-term securities. Empirical evidence is consistent with this rate’s being in the vicinity of 3% per annum (net of inflation). Direct evidence on time preferences for health outcomes is also consistent with a discount rate of 3%.

The panel therefore recommends the use of a real, 3% discount rate in the reference case. Before discounting, all costs should be adjusted for inflation. Because many published CEAs have used a discount rate of 5%, future analyses should include sensitivity analysis using 5% as well as other rates in the range of 0% to 7%. The discount rate should be reviewed and possibly revised periodically, to reflect important changes in economic conditions. To ensure that analyses will remain comparable, however, both 3% and 5% should continue to be used for at least the next 10 years.

### Handling Uncertainty in CEA

Cost-effectiveness analyses are subject to uncertainty with regard to estimates of effectiveness, the course of illness, HRQL consequences and preferences, and health care utilization and costs. Users of analysis need information on the degree to which CEA conclusions might change with changes in assumptions or values.

Sensitivity analysis is an appropriate tool with which to respond to this need.
The simplest method, which should be used in all CEAs, is a univariate (1-way) sensitivity analysis, in which estimates or assumptions are changed one at a time. These establish where uncertainty or lack of agreement about some key parameter's value could have substantial impact on the conclusion of a CEA. They suggest areas where efforts to obtain additional data might be warranted in terms of impact on decisions and, conversely, areas where additional precision would be unlikely to change results. For example, the estimated cost-effectiveness of thrombolytic therapy in older patients with suspected myocardial infarction was shown to be stable when the efficacy of therapy, the prevalence of myocardial infarction, or the incidence of stroke was varied, thus suggesting that further research to evaluate the risks and benefits of streptokinase in this age group was not warranted.

One-way sensitivity analyses underscore the overall uncertainty in the C/E ratio; therefore, analysts should also conduct multivariate (multway) sensitivity analyses, in which several parameters or assumptions are changed at the same time, for important parameters. If possible, a reasonable confidence interval or credible interval for the C/E ratio should be estimated based either on statistical methods or on simulation. The value of multivariate sensitivity analysis is greatest when there is reason to believe that estimation errors for key parameters are correlated, for example, if studies that overestimated the effectiveness of thrombolysis also tended to underestimate the associated risk of stroke. Several methods for performing such statistical analyses and simulations are described in the panel's full report.

**CONCLUSION**

The panel recognizes that many of the recommended methodological approaches are not broadly practiced at present, and there may be gaps in the availability of data to satisfy the criteria for the reference case. For example, there is no broad consensus on which health-state classification systems are suitable for CEA, and values of community-based weights for some systems are not publicly available. The ability to weight years of life by population averages of community preference weights is limited by the lack of appropriate (age- and sex-specific) population data for some systems. In the area of costs, there are no well-accepted methods for determining time costs for individuals outside the labor force, and few good-quality data on resource use, reflecting costs rather than charges and clearly applicable to the populations under study, are readily available to analysts.

The intention of these recommendations is to move the field of CEA closer to standardization in the near term where possible and to identify desirable practice where optimal methods are not currently feasible. All of these recommendations for the reference case, and for CEAs in general, are subject to a "rule of reason." When a parameter estimate or an element of the analysis is unlikely to have an appreciable effect on the result, then it may be acceptable to use shortcuts to obtain them; expert opinion may be used to assess them; or they may be excluded from the analysis altogether. Examples may include the weights assigned to short-term and mild impairments of HRQL, costs of unrelated health care in added years of life, or the incidence or costs of minor side effects of treatments. The rule of reason applies if the cost of obtaining more precise estimates of the parameter in question would exceed the value of achieving more precision in the final cost-effectiveness result. However, the burden is on the analyst to justify suboptimal methods of parameter estimation or exclusion of effects from an analysis.

If researchers endeavor to follow a standard set of methods in CEA and to obtain the required inputs for their studies, much will have been accomplished toward improving the utility of this form of analysis. It is hoped that the recommendations contained here will stimulate rapid progress toward availability of the necessary data and tools, so that the practice of CEA can soon become as established as many other forms of scientific inquiry.

The panel on Cost-Effectiveness in Health and Medicine membership and staff: Norman Daniels, PhD; Dennis G. Fyhrbaek, PhD; Alan M. Garber, MD, PhD; Marthe R. Gold, MD, MPH; David C. Hadorn, MD; Mark S. Kamlet, PhD; Joseph Lipscomb, PhD; Bryan R. Luce, PhD; Jeanne S. Mandelblatt, MD, MPH; Willard G. Manning, Jr, PhD; Donald L. Patrick, PhD; Louise B. Russell, PhD; Joanna E. Siegel, ScD; George W. Torrance, PhD; Milton C. Weinstein, PhD.

We thank Kristine I. McCoy, MPH, for coordination, research, and editorial assistance in association with this project.

**References**