OBJECTIVE. Before cost-effectiveness analysis (CEA) can fulfill its promise as a tool to guide health care allocation decisions, the method of incorporating societal values into CEA may need to be improved.

DESIGN. The study design was a declarative exposition of potential fallacies in the theoretical underpinnings of CEA. Two values held by many people—preferences for giving priority to severely ill patients and preferences to avoid discrimination against people who have limited treatment potential because of disability or chronic illness—that are not currently incorporated into CEA are discussed.

CONCLUSIONS. Traditional CEA, through the measurement of quality-adjusted life years (QALYs), is constrained because of a “QALY trap.” If, for example, saving the life of a person with paraplegia is equally valuable as saving the life of a person without paraplegia, then current QALY methods force us to conclude that curing paraplegia brings no benefit. Basing cost-effectiveness measurement on societal values rather than QALYs may allow us to better capture public rationing preferences, thereby escaping the QALY trap. CEA can accommodate a wider range of such societal values about fairness in its measurements by amending its methodology.

Key words: Cost-effectiveness analysis; quality-adjusted life-years; person tradeoff; rationing; disability; quality of life; ethics. (Med Care 2000;38:892–901)

Before CEA can fulfill its promise as a tool to guide health care allocation decisions, the way societal values are incorporated into CEA may
need to be improved. In this article, we explore 2 values held by many members of the general public that are not currently accounted for in CEA measurement and cannot be accounted for within the current CEA framework because of its emphasis on quality-adjusted life-year (QALY) maximization. We propose an alternative to QALY maximization—societal value measurement—that accounts for these 2 values and ultimately may enable CEA to more accurately reflect how the public wishes to set health care priorities.

**How CEA Currently Measures Value**

CEA estimates the number of dollars it takes to yield 1 unit of health outcome. A commonly used unit is the QALY, which takes into account both the quantity of life and the health-related quality of life (HRQOL) of patients receiving health care interventions. Consistent with the Panel on Cost-Effectiveness, we use the term HRQOL to connote the values assigned to various health states. Often, the HRQOL values used in CEA are elicited by asking the general public to answer standard gamble, time tradeoff, or rating scale questions. These HRQOL values are thought to possess interval properties, so equal changes in HRQOL imply equal benefits. For example, on an HRQOL scale ranging from 0, for health conditions as bad as death, to 1, for full health, an intervention that improves patients' average HRQOL from 0.2 to 0.4 produces half the benefit of an intervention that improves people from 0.2 to 0.6.

A QALY is a year of full health. A year of anything less than full health produces <1 QALY. A chronic illness, with an HRQOL of 0.5, produces 0.5 QALY per year, as does an intervention that improves HRQOL from 0.2 to 0.4 produces half the benefit of an intervention that improves people from 0.2 to 0.6.

Figure 1 shows the result of a CEA for a health care intervention that extends a patient's life for 1 year in a suboptimal state of health (with an HRQOL of 0.7 out of 1); the extra year of life begins 1 year after the patient received the intervention. Without incorporation of HRQOL and time discounting values, the intervention has a cost-effectiveness of $10,000 per life-year gained. After incorporating these values, the intervention has a cost-effectiveness of $14,712 per QALY.

As this example demonstrates, conventional CEA already accommodates values for HRQOL and time discounting in its measurements. In a similar fashion, it could potentially incorporate other values. Given ongoing criticism that CEA does not accurately capture public values for how to set health care priorities, we must look for ways of expanding value measurement in CEA.

**Two Values That May Merit Incorporation Into CEA**

Experts have proposed a large number of values that could be used to guide resource allocation decisions that are not currently incorporated into CEA. For example, QALY maximization emphasizes average benefits to populations and ignores the distribution of such benefits. Similarly, some have criticized CEA because it does not give enough priority to life-saving interventions. For example, if concerned only about QALY maximization, decision makers would be indifferent between providing a life-saving intervention that produces 100 QALYs and a non–life-saving intervention that produces 100 QALYs. However, life saving may have particular importance to some people, who are less concerned with the quality of life than with the quantity. In addition, some people think CEA should give priority to those age groups that are deemed more socially valuable. For example, Williams, for example, makes a “fair innings argument,” suggesting that elderly people do not deserve as much life-saving medical technology as young people.

Some criticisms of CEA could be dealt with relatively simply by modifying QALY measurement. For example, it is theoretically possible to generate “age-adjusted QALYs” much in the way that QALYs can be discounted for time delays in receiving benefits. Nevertheless, not all criticisms of QALYs can be handled within the QALY framework. To illustrate the limitations of the framework, we discuss 2 values currently ignored by CEA that may deserve incorporation into CEA: giving priority to severely ill patients over others.
and avoiding discrimination against people with limited treatment potential owing to disability or chronic illness. These 2 values can be incorporated into CEA only by abandoning the current QALY framework.

**Priority to Severely Ill Patients**

Some critics contend that CEA measures do not place enough priority on helping severely ill patients. The importance of helping the neediest is the basis of some important philosophical theories of justice and has been the subject of many debates about the allocation of scarce medical resources. Moreover, empirical evidence suggests that many people think severely ill patients deserve priority when we allocate scarce health care resources. For example, when asked whether scarce resources should be directed to a severely ill group of patients who can receive small treatment benefits or a moderately ill group who can receive substantial benefits, most Norwegians and Americans favored the severely ill patients. Other studies in Norway, the United States, Spain, and Australia showed that people give greater priority to treating severely ill patients than would be predicted from HRQOL measurements. For example, in one study, people estimated the HRQOL of moderate knee pain as being 0.94, a value that would imply that curing this knee condition (thereby bringing an HRQOL improvement of 0.06) is ~1/16th as valuable as saving a person’s life. Yet these people felt that >1,000 people would need to be cured of moderate knee pain to equal the value of saving 1 person’s life. People not only are concerned with the change in HRQOL brought by treatment but also want to give priority to patients with more severe illnesses.

**Avoiding Discrimination Against People With Limited Treatment Potential Owing to Disability or Chronic Illness**

If maximizing QALYs were the sole goal of health care spending, then it would be less valuable to save the lives of people with chronic disabilities, because it would bring fewer QALYs. For example, if paraplegia has an HRQOL of 0.8, then saving the life of a person with paraplegia yields only 0.8 QALYs per year (Figure 2). Yet people think that saving the lives of patients whose treatment potential is limited because of disabilities is often equally as valuable as saving other patients’ lives. Consequently, the emphasis on maximizing HRQOL in CEA conflicts with some people’s attitudes toward the importance of saving the lives of people with disabilities and chronic illnesses.

Some may question the relevance of assessing people’s preferences for avoiding discrimination against people with limited treatment potential resulting from preexisting illness or disability. After all, analysts do not calculate the cost-effectiveness of coronary artery bypass surgery in subpopulations on the basis of “paraplegic” or “nonparaplegic.” However, it is important that CEA be able to account for a life-saving intervention that is specific to a particular illness (eg, protease inhibitors for AIDS patients). Moreover, CEA experts have recently called on analysts to refine their analyses so that the calculation of treatment benefits takes account of people’s underlying health. For example, a CEA of coronary artery bypass surgery would have to account for the less than optimal HRQOL of patients who require this surgery. Given criticism that CEA undervalues the lives of people with disabilities or chronic illnesses, it is useful to determine whether CEA can be modified so that the value it places on the treatment potential.
of people with chronic health problems will more accurately reflect public sentiment.

**An Obstacle to Incorporating These 2 Values Into CEA: The QALY Trap**

Preference for giving priority to severely ill patients could easily be incorporated into CEA by rescaling HRQOL measures. For example, if we really thought that a change in HRQOL from 0.1 to 0.2 (on a 0 to 1 scale) was twice as important as a change from 0.7 to 0.8, then we could rescale HRQOL measures so that the former conditions were given a value of 0.1 and 0.3, respectively. In it, the HRQOL of any health condition determines not only the benefit of curing the condition but also the benefit of saving the life of someone with that condition (Figure 3). This forces us to decide whether we think that saving the life of a person with paraplegia is equally as important as saving anyone else’s life or instead whether we think that curing paraplegia brings some benefit.

But this rescaling would not allow us to incorporate the second value we have discussed, the preference people have for avoiding discrimination against people who have limited treatment potential. For example, suppose the public thinks that saving the lives of people with paraplegia is equally as important as saving the lives of people who can be returned to full health. CEA would accommodate this value by rating paraplegia as having the same HRQOL as full health (an HRQOL of 1.0). With this rescaling, saving the life of either group of patients will bring 1 QALY per patient. However, this rescaling would also force us to conclude that people with paraplegia have the same quality of life as people without paraplegia and thus that curing paraplegia would not improve HRQOL. (No HRQOL is gained by “improving” patients from an HRQOL of 1.0 to an HRQOL of 1.0.)

The QALY model has us trapped. In it, the HRQOL of any health condition determines not only the benefit of curing the condition but also the benefit of saving the life of someone with that condition (Figure 3). This forces us to decide whether we think that saving the life of a person with paraplegia is equally as important as saving anyone else’s life or instead whether we think that curing paraplegia brings some benefit.

The Panel on Cost-Effectiveness explicitly addressed the issue of whether CEA discriminates against people with disabilities. Many in the disability community are upset that some HRQOL ratings used in CEA are based on general public attitudes rather than direct measurements from patients. The disability community points out that the general public often underestimates the HRQOL of chronic illness and disability. The panel argued that even if the public underestimates the HRQOL of chronic illnesses, this would lead to an overestimation of the cost-effectiveness of interventions that prevent or treat these illnesses.

Although this argument may allay the concerns of the disability community, it is an unacceptable defense of erroneous measurement. When the resources available for health programs are limited, an overestimate of the cost-effectiveness of one program must necessarily reduce the relative cost-effectiveness of other programs. Moreover, the panel’s argument does not adequately address concerns that HRQOL estimates (whether from patients or the public) reduce the value of the lives of people who have chronic illnesses or disabili-
ties. No matter whose values are incorporated into CEA, according to the current model the benefit of saving the life of someone who has paraplegia is less than the benefit of saving the life of someone who can be returned to near-perfect health. And if public HRQOL estimates rather than patient estimates are used, the benefit of saving these patients’ lives would be given even less value. In trying to assuage fears that CEA discriminates against people with disabilities, the panel ignored the QALY trap. With the present methods for measuring cost-effectiveness, we cannot say that saving the life of a paraplegic is equal to saving the life of a person who can be returned to full health while simultaneously saying that there is still some benefit from curing paraplegia. If paraplegia has an HRQOL of 0.8 for the individual, then saving the life of a person with paraplegia for a year must be judged as having a societal value of 0.8 QALYs, and curing a person of paraplegia for a year must bring 0.2 QALYs.

How Societal Value Measurement Allows Us to Escape the QALY Trap

The QALY trap exists because HRQOL measurement is being asked to capture not only patients’ quality of life but also the societal value of treating patients with various health conditions. When HRQOL measures are supplemented with separate societal value measures, we will no longer be caught in the QALY trap.

By societal value, we refer to the strength of public preferences for giving priority to various competing health care programs. If people think, all else equal, that severely ill patients should receive priority over less severely ill patients, then programs directed toward severely ill patients have greater societal value. Some of us have proposed that societal value could potentially be measured by the person tradeoff (PTO) method. In this method of eliciting values, people might be asked how many patients need to be cured of paraplegia (program C from Figure 3) to be equally as important as curing 100 people of a life-threatening illness and returning them to full health (program A).

Suppose the societal value of program A is given an arbitrary value of 1.0. Now, suppose people think that program B, which saves the lives of people with paraplegia, should receive the same priority for funding as program A. In other words, suppose people think that it is equally beneficial to save 100 lives with program B and 100 lives with program A. Then, program B must also have a societal value of 1.0. Now suppose the same people think that curing 600 patients of paraplegia (program C) is equally as important as saving 100 otherwise healthy people’s lives (program A). Program C therefore has one-sixth the value of program A and consequently a societal value of 0.16.
Figure 4 shows the hypothetical societal values of programs A, B, and C and of a series of other programs.

What is the benefit of separating societal value from HRQOL measurement? As shown in Figure 3, the HRQOL brought by programs B and C (in conventional CEA) must sum to the number of QALYs brought by program A (1 QALY). However, as shown in Figure 4, the societal value of programs B and C add to 1.16. We have escaped the QALY trap. We can now say that saving the lives of people with paraplegia is equally as valuable as saving other people’s lives while still acknowledging that it is beneficial to cure people of paraplegia.

By incorporating societal value measures into CEA, we can account for the priorities that the general public places on helping severely ill patients. For example, in Figure 4, health care interventions that improve people’s average HRQOL from 0.2 to 0.8 (an improvement that brings 0.6 QALYs per year) are given a hypothetical societal value of 0.7 to illustrate how societal value measurement can give extra value to treatments directed at helping severely ill or disabled patients. In addition, societal value measurement can account for public attitudes toward the importance of helping patients who have limited treatment potential because of pre-existing disability or illness. In Figure 4, we present hypothetical societal value data in which the value of saving the life of a person with paraplegia is equal to that of saving someone who could be returned to perfect health, despite the reduced HRQOL of paraplegia. We also present hypothetical data for life-saving treatments that affect people with more severe disabilities, with HRQOL of 0.2 and 0.5, respectively; in these 2 cases, we specified societal values of <1 to highlight an important point: Societal value measurement does not force us to conclude that saving the life of a person with any disability is equally as valuable as saving anyone else’s life.

Potential Criticisms of Our Proposal

Escaping the QALY Trap Creates Inconsistencies

Some may worry that the societal value approach to QALYs is inconsistent because it allows the value of programs like B and C to sum to a value >1. However, if society places the same value on saving the life of a paraplegic and saving the life of a nonparaplegic and if society also values the cure of paraplegia, then our suggestion is consistent with societal values. When we think about saving the life of a person with paraplegia, we are making a separate decision than when we think about curing paraplegia. When we decide about the benefit of curing paraplegia, our decision is not dictated by the value we think should be placed on saving the life of a person with paraplegia.

Societal Value Measurement Ignores Important Issues of Distributive Justice

Health care allocation decisions, which provide similar benefits and harms to people in a popula-
tion, are sometimes viewed as being more equitable, and thus more socially valuable, than allocation decisions that benefit some people much more than others.\textsuperscript{6,10,19,42} Policy experts have even developed mathematical measures of equity so that the relative equity of policy alternatives can be estimated.\textsuperscript{14,43–47} Measuring distributional equity in this way is a promising method of improving certain public policy decisions. For example, in deciding about a new environmental regulation, a regulation that causes tiny detriments in many people’s health may be preferred to one that causes extreme health problems in a small number of people.

Nevertheless, equity measures like these will not transfer easily into health care CEA. When a major public policy decision about an environmental law is being made, it is reasonable to conduct an extensive, one-time economic analysis that looks at equity issues. In contrast, CEAs are intended to guide a whole range of medical decisions, from a single hospital’s purchasing decisions to federal government decisions about reimbursement for new technologies. In one context, a health care system may face budget limits, forcing it to choose between offering the best available treatment to a portion of its patients or a less effective treatment to everyone. Another health care system, with fewer budget constraints, may be able to offer more of the best available treatments to everyone. As this example demonstrates, it makes little sense to calculate “distributional equity coefficients” for specific health care interventions. To the extent that these distributional preferences deserve a role in resource allocation, even CEA modified in the ways we propose will remain only a guide to resource allocation decisions.

\textbf{Methodological Questions About Societal Value Measurement}

Earlier, we suggested that societal values could be measured by the PTO method. This method can be best understood by contrasting it with measures traditionally used in CEA. Some have argued that the most theoretically pure HRQOL measure is the standard gamble,\textsuperscript{7} which asks people to state the highest chance of death they would take to rid themselves of a health condition. For example, if people are willing to take, on average, a 20\% chance of death to rid themselves of paraplegia, then paraplegia has an HRQOL of 0.8. Another common measure is the time tradeoff method,\textsuperscript{48} in which people are asked how much of their life span they would give up to improve their health. If people, on average, are willing to give up 20\% of their life span to be cured of paraplegia, it has an HRQOL of 0.8. Common to both the standard gamble and time tradeoff methods is an emphasis on people’s individual preferences for various health states–they think about the risk of death they would take (or amount of their life span they would give up) to rid themselves of a health condition. In theory, these individual preferences can be aggregated to provide societal estimates of a health state value, although such aggregation is controversial.\textsuperscript{49} More importantly, these preference elicitation methods lack a societal context for evaluating health states. They do not ask people to think about how society ought to decide between competing health care programs. Because CEA is supposed to help guide such decisions, the PTO method potentially elicits people’s values in ways more relevant to the task at hand.

Many methodological issues need to be worked out before we can confidently claim that the PTO method or other measures capture societal values with enough accuracy to include them in economic analyses. For example, societal value measures, like many such measures, are susceptible to framing effects\textsuperscript{50} and cognitive bias.\textsuperscript{27} Furthermore, some compromises may need to be made to incorporate societal values into complex CEA models. Markov models, for example, estimate the value of interventions affecting patients over time, where patients may have perfect health 1 year, mild congestive heart failure the following year, and severe congestive heart failure the next. It may be difficult to measure the societal value for each transition state in such a model.

Some might wonder if it will ever be possible to have an adequate measure of societal values, because people do not have well-established ideas of how they value different types of treatment programs and because the measurement of societal values is quite complex. Nevertheless, a perfect representation of societal values is not needed before they can be incorporated into CEAs. After all, few experts think that present HRQOL measures perfectly capture the quality of life of all health conditions. Yet most agree that imperfect HRQOL measures are preferable to measuring cost-effectiveness without taking any account of quality of life. Thus, with enough empirical work,
it should be possible to approximate societal values for treatment programs with enough accuracy to include them in CEA.

Societal Value Measurement Ignores Benefits of Interventions That Simultaneously Save Lives and Cure Disability

Imagine 2 life-saving treatments. Treatment A saves the life of a patient with a maximum HRQOL of 0.8 because of a preexisting disability. Treatment B simultaneously saves the patient’s life and cures her of a similarly severe disability. In the QALY model, treatment A would yield 0.8 QALY, and treatment B would yield 1 QALY. Yet in our model, both would have a societal value of 1. Critics of our proposal may argue that the QALY model is able to recognize that treatment B brings more benefit than treatment A, whereas our model can not.

However, this is simply the QALY trap in reverse. The greater ability of the QALY model to distinguish between these 2 treatments occurs, as we have discussed, only because the QALY model undervalues program A. For example, if 2 patients were near death and treatment A was feasible for one while treatment B was feasible for the other, many people would think that these 2 patients deserve equal treatment priority. Moreover, those who favor the latter patient would do so much less than would be suggested by their relative QALY gains. The first patient will not be seen as being 80% as important to treat as the latter. If this were a single patient with 2 alternative treatments, one of which would cure the disability and the other which would not, it is obvious which treatment we should choose. But most health program decisions concern allocations among different groups of patients. In such contexts, the QALY model’s undervaluing of life saving treatments affecting people with chronic illnesses or disabilities is a bigger problem than our model’s inability to distinguish between treatments that simultaneously save lives and cure previously incurable chronic illnesses or disabilities.

Society’s Values May Be Unjust

Our proposal emphasizes ways to make CEA more closely capture public allocation preferences. But public preferences themselves may be unjust.

Fixing CEA: Incremental Repairs Would Only Make Things Worse

Societal value measurement will never be able to account for all values relevant to how society ought to allocate health care funds. This incompleteness raises the possibility that a modified CEA taking account of societal value would create the illusion of completeness: People would see that CEA accounted for important values and forget that it still fails to account for other important values. In contrast, it could be argued that CEA in its current form is so blatantly inadequate that it never fools anyone into thinking it is a perfect allocation tool. We are unaware, however, of any evidence that suggests that this is true. Moreover, the above argument could equally be applied to CEA as it is currently measured. That is, CEA might already make things worse by producing a beguiling set of numbers to assist with economic evaluation decisions. More generally, this argument could be used against virtually all economic evaluations, which usually fall somewhat short of gold standard measures.

The hazards of improving CEA measurement, however, must be recognized. Current CEA studies are already influencing allocation decisions. For example, CEA is being used by some government and third-party payers to determine reimbursement for new medical technologies. Given the growing use of CEA to guide health care decisions, we think it is important to do whatever we can to improve CEA so that it more closely (even if...
imperfectly) captures public allocation preferences. At the same time, decision makers need to be better informed about the strengths and weaknesses of CEA.

Conclusions

Health care resource allocation decisions are fraught with complexities, so no measurement tool will capture every relevant value. Our goal should be to capture the most important values that will affect the most people. Two values captured by our societal value model come closer to approximating public rationing preferences than the current QALY model. Moreover, the modifications we propose are well within the spirit of how CEA has already tried to capture other values. If CEA is ever to be accepted by advocates for people with disabilities or by the general public, a better attempt needs to be made to find out what the public really values in health care programs beyond simply maximizing QALYs.

Acknowledgments

We gratefully acknowledge Ellen Wise for assistance in manuscript preparation and Andrea Gurmankin and Allan Detsky for comments on an earlier draft.

References

27. Ubel PA, Scanlon D, Loewenstein G, Kamlet M. Individual utilities are inconsistent with rationing


