Principles of Cost-Effectiveness Research

Theodore G. Ganiats, MD, and Lawrence J. Schneiderman, MD
La Jolla, California

Recent advances in the theoretical concepts behind cost-effectiveness evaluation have led to a marked improvement of the techniques used in cost-effectiveness research. Any effective program will have not only dollar costs and dollar benefits but also nondollar costs and benefits. Cost studies evaluate the relative weights of these costs and benefits to determine whether the program's value is worth the expense. Key elements to consider include direct costs of the program, costs of future workup or treatment, costs of evaluating and treating complications resulting from the program, future medical costs averted, discounting of future costs and health effects, and a sensitivity analysis. Despite the recent theoretical advances, cost-effectiveness research raises multiple controversies, including some key ethical issues.

Because of the increasing demand for and cost of health care resources, there has been much concern about the need for improved cost-effective measures in clinical practice. These thoughts have spawned a host of cost studies, and family physicians are becoming increasingly involved in such research. This research can be used to evaluate programs dealing with both the diagnostic and therapeutic aspects of patient care.

One rationale for cost-effectiveness research is policy formulation.[1] Such questions as, Should mammograms be done? Is the routine admission chest x-ray examination worthwhile? and Should patients with a pharyngitis receive a throat culture? all reflect policy issues; the answers influence what the American Cancer Society recommends or what an insurance company will reimburse. Other questions (eg, Should a patient with an acute exacerbation of chronic bronchitis receive antibiotics?) deal more directly with patient treatment issues but also are cost-effectiveness questions, as cost-effective management often is the best management. Cost studies can also assist in the determination of the appropriate cutoff values for normal and abnormal by evaluating the optimal sensitivity and specificity of a test.

This article reviews basic concepts of cost-effectiveness evaluation and recent advances in cost-effectiveness research. It is divided into four parts: (1) a definition of key terms used in cost studies, (2) a review of the three basic types of cost studies and the relative strengths and weaknesses of each, (3) key points for consideration in both the reader's own research and the evaluation of studies performed by others, and (4) a discussion of ethical issues implied by cost-effectiveness research.

DEFINITION OF TERMS

Costs and Benefits

Any program will have not only dollar costs and dollar benefits but also nondollar (usually health) costs and benefits. The terms *cost* and *benefit* have lost clarity, as they have been used to refer to both the dollar and the nondollar elements of the analysis. To help eliminate the confusion, costs will refer to dollar expenditures, and benefits will refer to dollar savings.

Effects (or Health Effects)

The effects are the nondollar elements of a program. Positive effects are those desired, the various program outcomes that measure improved health. Negative effects are the untoward health outcomes of the program, usually considered the negative side effects that increase morbidity or mortality.

Perspective

There are many ways of interpreting the phrase *cost effective:* the program either (1) saves money, (2) is effective,

Submitted, revised, April 29, 1988.

From the Division of Family Medicine and the Division of Health Care Sciences, Department of Community and Family Medicine, University of California, San Diego, School of Medicine, La Jolla, California. Presented at the 21st Annual Spring Conference, Society of Teachers of Family Medicine, Baltimore, April 27, 1988. Requests for reprints should be addressed to Dr. Theodore G. Ganiats, Division of Family Medicine T-007, UCSD School of Medicine, La Jolla, CA 92093.

(3) saves money and is at least as effective as the standard, or (4) has an additional effectiveness worth the additional cost.[2] The study may be performed from the perspective of the individual, institution, or society. The perspective chosen is key in determining the correct interpretation of data, and for this reason, clarifying the perspective is critical. Since cost studies usually answer questions of social policy and deal with a limited commodity, the societal perspective is generally considered the most appropriate.[3,4]

Many cost studies evaluate only the direct costs and benefits of a program. Given the societal perspective, the analysis takes on a broader meaning, and it becomes clear that all costs and benefits referable to the program should be counted. Sample costs and benefits are listed in Table 1 along with an example of each type of element given a hypothetical evaluation of amniocentesis for Down's syndrome. Details are discussed below.

TABLE 1. SAMPLE COSTS AND BENEFITS FROM A SOCIETAL PERSPECTIVE

Example: Genetic amniocentesis for Down's syndrome

Direct costs	Amniocentesis, physician fees, pathology
Future workup and treatment	Repeat of equivocal test, therapeutic abortion
Evaluation and treatment of program complications	False-positive and false-negative, fetal loss, infection, bleeding
Future costs averted	Surgeries and other medical care for infant, special education
*Future medical costs	Prenatal care for another pregnancy (potential)
*Future salary savings	Increased salary potential for someone without Down's syndrome

** Areas of current controversy (see text for description and references)*

TYPES OF COST STUDIES

As mentioned above, a program has dollar costs, dollar benefits, and nondollar health effects. Cost studies evaluate the relative weights of these costs and effects to determine whether the program's value justifies the expense. This evaluation can be done by examining only the monetary aspects of the program (only dollar costs and benefits) or by also examining nondollar health effects. The types of cost studies are shown in Table 2. A *cost study* refers to any cost evaluation. *Benefit – cost analysis* assigns a dollar value to the dollar costs and the dollar benefits and may also try to assign a dollar value to the nonmonetary outcomes. *Cost/effectiveness analysis* considers the ratio of the total net dollar cost (numerator) to a net nondollar effect (denominator). A *benefit–cost/utility analysis*[5–7] is a specific type of cost/effectiveness analysis in which the denominator is a standardized utility function, that is, a standardized scale that expresses the patient's relative preference for the various outcomes.

TABLE 2. TYPES OF COST STUDIES

Benefit – cost analysis
- Calculates the difference between the program's benefits and costs
- Usually reported in dollars

Cost/effectiveness analysis
- Calculates the ratio of the program's net dollar cost to a net nondollar effect
- *Examples:*
  - Cost/effectiveness analysis: Dollars per cancer found
  - Benefit – cost/utility analysis: Dollars per well-year

Benefit–Cost Analysis

The most basic form of cost study is the benefit – cost analysis. In the past this was called a cost – benefit analysis, which implies subtracting the value of the benefits from the value of the costs. In the ideal situation, in which benefits are greater than costs, using this definition results in a negative number. To avoid this discrepancy of having a negative number indicate a positive outcome, many researchers now subtract the cost value from the benefit value, so that a positive result indicates a positive evaluation.[8]

Net benefits are the savings in medical and nonmedical costs. For example, preventing a Down's syndrome birth (Table 1) saves medical costs because numerous surgeries, hospitalizations, and physical therapy treatments are avoided. Nonmedical savings derive from the prevented institutionalization or special education programs that will not be needed. It is important to note that net benefit does not imply total benefit. The amount of money saved by preventing one year of special education, for instance, is equal to the cost of that special education program minus the cost of a normal education. In all cases (benefits, costs, and health effects) the program participants must be compared with an appropriate control group.

Net costs include the total direct costs of the program plus the costs of side effects or complications of the program. In the Down's syndrome example program costs include those of the amniocentesis, tissue cultures, and patient follow-up. Any costs of treatment for infection, miscarriage, or bleeding induced by the amniocentesis are also considered. Future medical costs for prolonged life and future income salary generated are controversial areas discussed below.

As stated above, the benefit – cost analysis involves only dollar amounts. Physicians often argue against ignoring the nonmonetary aspects of patient care. Some benefit – cost studies approach this argument by assigning a dollar value to the nonmonetary effects. To do so is difficult at best (What is the dollar value of vaginal discharge for two weeks?) and introduces one of the major drawbacks of this approach.

A benefit – cost analysis evaluating only the financial elements of a program is based on the assumption that the goal of health care is to save dollars. Clearly, health care goals also include nondollar outcomes that should be evaluated. As described by the World Health Organization, the goals of health care are to add years to life and life to years,[9] or in other words, to increase the quantity and quality of life. To convert all health effects to dollars illustrates a major deficit of the benefit – cost analysis: it is difficult and perhaps inappropriate, and if one looks only at the direct financial aspects of the program, major factors may be missed.

A second major drawback of benefit – cost analysis involves opportunity costs: society prefers to support the most efficient programs, as resources used by one program cannot be used at the same time by another. Take, for example, programs A and B, both of which cost the same (ie, the dollar outcomes of benefit – cost analysis are identical). Assume program B has greater health effects (more positive nondollar outcomes). Even though program A may be equally cost effective using a benefit – cost analysis, program B makes better use of the same dollars and would therefore be the preferred program. This criticism of benefit – cost analysis can be addressed by the cost/effectiveness analysis.

Cost/Effectiveness Analysis

A cost/effectiveness analysis divides the results of a benefit – cost analysis by a nondollar program outcome measure. The preceding statements related to the calculation of the net dollar costs therefore also hold for cost/effectiveness analysis. The benefit – cost analysis problems with opportunity costs are often addressed because nonmonetary outcomes can be more directly assessed. A major criticism of the cost/effectiveness analysis model is the nonstandardized denominator (or program outcome). Recent examples from *The Journal of Family Practice* include positive electrocardiogram,[10] cancers and polyps found,[11] and year of life saved.[12] These studies are sound, but since the denominator varies, it is impossible to compare the relative effectiveness of the different programs. For example, which is more cost effective: spending $245 per positive electrocardiogram, $5,000 per cancer or polyp found, or $24,000 per year of life saved? Without knowing (1) what a positive electrocardiogram or cancer or polyp found means in terms of treatable pathology, (2) the costs of that treatment, and (3) what, exactly, a year of life saved means, the reader is unable to choose among the three.

Even if the denominators are the same, they may not be equivalent. Take the cancer or polyp found example. Is the program that costs $5,000 per cancer or polyp found better than the program that costs $15,000 per cancer or polyp found if the latter finds a greater proportion of polyps and thereby prevents more cancers? Which program is better if in the more expensive program all cancers found are Dukes' stage A, but in the less expensive program all are Dukes' stage C? To compare different studies, the denominator must be standardized and equivalent.

Benefit–Cost/Utility Analysis

The denominator (outcome measure) should reflect the goals of the program, and as stated above, the goals of medicine are to improve the quality and quantity of life. The denominator in a cost analysis should therefore be an outcome measure reflecting both wellness and time.[3,4,8] The well-year (WY)[6] and quality-adjusted life year (QALY)[8] are two such measures. A well-year is defined as one person living a healthy life for one year and can be calculated by multiplying the person's wellness (described below) and time.* Thus ten people living five years at 70 percent wellness ($10 \times 5 \times 0.7 = 35$ WY) is the same as one person living 35 years at 100 percent wellness. Since both the well-year and the quality-adjusted life year are utility measurements, some authors have named an analysis that evaluates cost effectiveness using this standard denominator a cost/utility or benefit – cost/utility analysis.[23,24] This latter term clarifies the analysis: divide the result of the benefit – cost analysis (the net cost) by the net outcome (utility) produced by the program.

The benefit – cost/utility analysis (1) addresses appropriate health effects (quantity and quality), (2) allows for a comparison of quite different programs because the numerator and denominator are in standard units, and therefore (3) addresses issues of opportunity costs. For these reasons the benefit – cost/utility analysis is preferred by many for health policy formulation and program evaluation.

But when is a program cost effective? Who determines what the cutoff for cost effective will be? This answer is perhaps the most pervasive misconception regarding cost studies. Except for programs that save dollars and have

** The well-year is a health status index based on empiric measures of a large population. Its validity and reliability have been repeatedly demonstrated,[7,13] and the Well-Year has been used in a broad range of studies evaluating cost effectiveness,[14–19] resource allocation,[20,21] medical care quality,[22] community health status and program analysis,[20] and others. A more detailed description is available from Kaplan et al.[6]*

positive health effects, no program should be simply declared cost effective at the whim of an investigator. All cost-effective programs are relatively cost effective; they are cost effective compared with some arbitrary standard or previously accepted program. Who sets the arbitrary standard? Given the societal perspective, society should set the standards for cost effectiveness. Assuming that programs society supports are cost effective, and those programs society does not support are not cost effective, Kaplan et al proposed that society deems all programs costing under $20,000/WY cost effective and those costing more than $100,000/WY not cost effective (1980 dollars).[6] (For those programs costing between $20,000 and $100,000/WY, society is not yet consistent). Sample benefit – cost/utility ratios are in Table 3.

TABLE 3. BENEFIT – COST/UTILITY OF VARIOUS PROGRAMS*

Program	Benefit – Cost/Utility	Year
Phenylketonuria screening[15]	$ 2,900/WY**	1973
Thyroid screening[16]	$ 3,600/WY	1981
Coronary artery bypass (three-vessel)[18]	$ 7,200/WY	1985
Estrogen therapy for menopausal symptoms[19]	$18,600/WY	1980

** Since inflation produces variations in the value of the dollar, the year the study is performed should accompany any statement of relative cost effectiveness. In addition, in these studies differences in preference weights and discount rates may make direct comparisons inappropriate*

*** Well-year*

OTHER CONSIDERATIONS

Discounting

The costs and outcomes of a program rarely occur at the same time; there are both present and future costs and effects. For example, a screening program results in dollar costs and negative health effects in the present, while most dollar savings and positive health effects occur in the future. Most researchers apply a discount rate to future costs. One rationale for discounting is investment potential. To claim that $100 saved ten years from now is the same as $100 saved today ignores the investment potential of today's dollar. Another rationale for discounting is consumer preference; when asked, most people would prefer goods and services a dollar can buy today rather than waiting. Given this preference, one dollar today is valued more than a future dollar, and the latter must be discounted. It is important to emphasize that inflation is not a reason for discounting future costs, that discounting should occur after adjusting for inflation.[3]

Applying a discount rate to future health effects at first seems unjustified, for how can one year of life in the future be worth any less than one year of life today? Discounting future health effects never intends to decrease the value of future life, but it does assure the relationship between dollars and wellness remains constant. If future dollars are discounted (as they must be), the future health effects must similarly be discounted or the value of future life increases relative to the discounted dollar. There is some debate, however, whether future costs and future effects should be discounted at the same rate. If future health is likely to be better than current health (that is, the health of the nation is improving), perhaps future health effects should be discounted at a rate lower than costs. The converse holds if the prospects for future health are less than for current health.[25]

The whole field of discounting often raises more questions than it provides answers. Though various discount rates (from 3 to 10 percent) have been used in the past, the exact number is not so critical as the concept. To assist in the comparison of different studies or programs, standard discount rates of 0 percent and 5 percent have been suggested.[4] Further discussion of the issues related to discounting can be found by Russell,[4] Weinstein and colleagues,[3,8] Fuchs and Zeckhauser,[25] and Keeler and Cretin.[26]

Future Medical Costs and Salary Savings

Two controversies dominate the calculation of the net cost of a program. First, how should future medical costs resulting from life prolonged by a successful intervention be valued? For example, how should the future medical costs resulting from a prevented death from myocardial infarction be counted? Since these future costs would not have occurred had the person died, some feel these additional costs should be added to the total costs of the program. Others argue that society accepts this increase in health care costs, that after the life is saved, the slate is clean and the analysis should end.

Clearly, there are two issues. Few would argue that after a child's life is saved by a pertussis vaccine, all future medical costs for the child should be added to the calculation. The program did fulfill its purpose, a life was saved, and "normal" future medical expenses (including the results of the normal aging process) are to be expected but not counted. On the other hand, should a death from myocardial infarction be averted, but the patient require years of intensive medical and surgical treatment for heart disease, the future medical costs should be counted. In this latter case the medical costs are a direct result of the same disease process originally addressed. Future medical costs continue to inspire intense debate.

The second area of controversy is future salary savings. Assume a 50-year-old is found to have Dukes' stage B colon cancer and is cured instead of dying in three years.

This person will now be able to work 15 years (until 65 years) instead of three years, and this 12-year difference of productive life, some argue, should be counted as a financial benefit, since it increases the resource pool from a societal perspective. Others argue that future salary savings are already included in the preference weights for wellness, and to count these savings again is double dipping. The debate on this issue also continues.

Retrospective vs Prospective Analysis

Cost analysis may be either retrospective or prospective. At first glance a retrospective analysis appears ideal: it is quicker and less expensive than a prospective study, all the important financial data are usually available from a computer database, and patient outcome is known. Some retrospective studies rely on a literature review to assess a broad experience, while for others the data come from a retrospective look at one program.

Literature reviews have difficulties, since the studies reviewed often evaluate slightly different questions in different populations at different times, and averaging or summarizing results from these studies can force complicated statistical problems. One solution is meta-analysis, a technique that was recently reviewed.[27] While the retrospective evaluation of one program avoids this problem, there is another difficulty that neither form of retrospective analysis addresses.

While a retrospective study may accurately quantify costs, health effects can be only estimated in a retrospective study: while the number of patients with a complication may be known, exactly how each patient reacted to the complication is unknown. Merely knowing that headache occurred in 10 percent of the patients, for example, does not address the degree to which the headache affected the patients. This information is vital in calculating the true net health effect of the program.

Prospective research, then, is the preferred method for cost analysis when time and finances allow. Retrospective cost analyses should be considered as cautiously as other retrospective studies and are best used as a preliminary analysis to justify the expense of prospective research.

Comparison Standard

As discussed above, when evaluating the cost effectiveness of a program, a basic question is, this program is cost effective compared to what? The issue is somewhat complex, for it depends on how many programs are to be evaluated.

Single Program

Single-program evaluations evaluate, Is program X cost effective? and the comparison standard is usually no intervention. For example, is fecal occult blood screening for colon cancer more cost effective than not screening? In this case the sensitivity (or the false-negative rate) usually becomes unimportant, since the patient with a false-negative result merely simulates the control group. A false-negative result on fecal occult blood screening examination simply means that the patient was treated as though he were never screened: no benefit accrues, but the patient is not placed at any special risk (except the risk of false reassurance), and the screening cost is negligible. A false-positive result incurs the negligible dollar costs of the initial screening plus the workup costs (often $1,000) plus the negative health effects of complications of the workup. The false-negative result fails to help the analysis; the false-positive hurts the analysis.

Multiple Program

Weinstein[28] has argued that cost effectiveness should always be assessed by comparing a procedure with less costly alternatives rather than with just the option of doing nothing. Evaluating two programs, eg, fecal occult blood screening and flexible sigmoidoscopy, may be done by comparing both to no intervention or by comparing both to a mythical "gold standard" with 100 percent sensitivity and specificity. In the later case, both the false-positive and false-negative rates of each program will be important. Fecal occult blood screening is much less sensitive than flexible sigmoidoscopy. Since the sensitivity is relatively unimportant in the analysis comparing a program to no intervention, but important in the analysis comparing a program to a mythical gold standard, the comparison of fecal occult blood screening and sigmoidoscopy may change depending on the perspective of the study. Early clarification of this perspective and its limitations is therefore important.

Another consideration involves procedures that are not mutually exclusive. In the above example the result of the fecal occult blood screening affects the result of the sigmoidoscopy, and vice versa. In other words, the findings on flexible sigmoidoscopy are more likely to be negative after the patient has a negative screening by fecal occult blood screening than before screening. When two tests are not mutually exclusive, the proper sequencing of the tests becomes important. Which test is the more efficient can be determined through a cost study, and the sequencing of tests can be evaluated by decision analysis and an evaluation of the marginal costs.[3]

Sensitivity Analysis

Many of the numbers used in a cost study are estimates. Some figures (eg, hospital costs) may be known in one setting but vary significantly in different parts of the country. Other figures (eg, life expectancy in a patient

with a metastatic cancer) may vary considerably, and only averages are known. Still other figures (eg, proportion of pyelonephritis in uncircumcised children that could be prevented by circumcision) can at best be only estimated by expert opinion, which itself will vary. It is clear that, depending on whether high or low estimates are chosen, the end results of an analysis may dramatically change.

For this reason a sensitivity analysis should be performed on all critical elements of the analysis. For example, if the hospital cost for a disease X ranges from $2,000 to $6,000, the benefit − cost/utility ratio should be calculated using both numbers. For many factors the relative importance of the discrepancy is minimal, and the sensitivity analysis demonstrates the strength of the conclusions of the study. When the sensitivity analysis does affect the analysis critically, the reader may be able to apply the number most appropriate to his own setting in deciding what is best. Sensitivity analysis thus illustrates critical areas for research by highlighting the program's key variables.

On occasion a sensitivity analysis may be avoided by assuming a best-case or worst-case scenario from the start. If, for example, routine tonsillectomy is felt to be not cost effective, all assumptions in the analysis should favor tonsillectomy. If the procedure fails to be cost effective even in this best case, a sensitivity analysis is unnecessary.

Policy Space

Figure 1 shows a two-dimensional health policy space where net costs are plotted on the X axis and wellness is plotted on the Y axis. Programs that fall in the right upper quadrant are all cost effective: they save both money and generate wellness. The left lower quadrant contains programs that lose both money and health, clearly programs that should be avoided; the right lower quadrant contains programs that save money at the cost of health. Most programs in medicine plot in the left upper quadrant, costing money but promoting health.

For programs falling in the left upper quadrant, benefit − cost/utility ratios are appropriate, and the −$20,000/WY mentioned earlier may be used as a marker for cost effectiveness. The more cost effective the program, the smaller the ratio (that is, the number will be less negative). Ratios are not helpful for those programs in the upper right quadrant, and the results should be reported as, for example, "saves $12,000 and 10 WY." The following example demonstrates the rationale.

Take two very cost-effective programs, program C and program D. Both programs are so cost effective they actually save money and save well-years (that is, in the long run they cost less than doing nothing, and they improve life). An example of such a program may be the Hemophilus influenzae type b vaccine.[29] Also assume program C saves $10,000 and increases life by 5 WY, and that program D saves $12,000 and increases life by 10 WY. Merely calculating ratios shows program C is $10,000/5 WY = $2,000/WY, and program D is $12,000/10 WY = $1,200/WY. Since the ratio for program C is greater (more positive) than that for program D, program C seems the better program. Closer evaluation reveals the faulty logic: program D must be the better program, for it offers much more (saving $2,000 more and gaining 5 extra well-years) than program C. Obviously, ratios are not appropriate when both the numerator and denominator are positive (or negative).

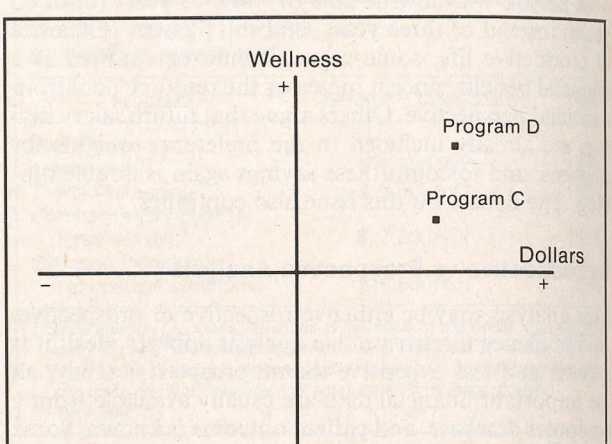

Figure 1. Two-dimensional health policy space

Another example illustrates one more difficulty that occurs when both costs and effects are positive. Which is better, the program that saves $2,000 and 12 WY or the program that saves $8,000 and 4 WY? Ratios are not appropriate (as illustrated above), and yet the answer is not obvious. The analysis of marginal costs in calculations in the right upper quadrant of policy space, along with a more detailed description of policy space, was recently reviewed by Anderson et al.[23]

ETHICS OF COST EFFECTIVENESS

Cost studies, because they provide quantitative measures of complex qualities, have proved to be extremely seductive to unwary health policy planners making rationing and resource allocation decisions. This application of cost/effectiveness analysis has raised several philosophical[30] and ethical concerns.

For example, recommendations on upper limits with respect to cost-effective therapeutic programs described in this paper are based on the underlying assumption that

the present health budget is fixed and universally accepted. At the present time the United States devotes approximately 11 percent of its gross national product to health care. This percentage has approximately doubled over the past few decades, however, it is not clear that society has established limits as to what it is willing to spend in this area. Obviously, if the health budget were to continue to enlarge, more expensive therapeutic programs might be considered cost effective.

There are at least two levels at which rationing and resource allocation decisions are made. At the first level, the health budget competes against all other societal expenditures, including those for military, education, public works, etc. At the second level, society must then decide what to do with what has been designated as the health budget, eg, how much should be spent on prenatal care, heart transplantation, renal disease, etc.

Another assumption made when a cost study is used as the basis for rationing and resource allocation is that such decisions should be prioritized by cost, that is, cheaper therapeutic programs should be favored over more costly therapeutic programs. This prioritization is itself a qualitative decision. It is conceivable that other values might be judged more important by a society. For example, society might favor newborns over the elderly, veterans over nonveterans, and socially productive individuals over their less productive counterparts. In the last case, for example, society might accept costs of $100,000 per year devoted to an individual who is an important government figure, yet not for an individual who is retired and confined to a convalescent facility, although both might regard their quality of life equally.

A contemporary philosopher, John Rawls,[31] has proposed that any inequality of social benefits is acceptable only when such inequality benefits the least fortunate. Thus, for example, a patient in a nursing home might accept the preferential treatment of a highly influential and important government figure if that government figure made the world safer for all people including those in nursing homes.

Society, on the other hand, might reject any deliberate favoritism and seek to provide equal health care for all. Even so, cost studies put the elderly, the poor, minorities, and women at a disadvantage because of social economic characteristics that may be completely unrelated to the medical therapy under scrutiny: the elderly will be more likely to die of other diseases, the poor and minorities (usually) will require more ancillary support such as transportation and social services, and women obtain lower wages for their work. Thus, these groups will tend to be disadvantaged when it comes to cost study outcome measures.

Another ethical concern about cost studies is that death is arbitrarily measured as zero, the lowest measure of quality of life. Yet, ethicists are aware that many patients feel that some qualities of life are worse than death—a perception that has given rise to social forces leading to statutes providing for living wills, durable powers of attorney, and other legal means of refusing medical treatment even in the face of shortening life. So far, few cost-effectiveness calculations have taken this issue into account.

Ethicists are also concerned that the data upon which patient preference indicators are based have been obtained from healthy persons who are asked to imagine how they would feel if they were ill. Whether such projections are accurate has ethical consequences, as cost-effectiveness calculations might well misrepresent true patient wishes at the time of illness.

Cost studies also equate degrees of improvement without regard to functional level. Thus, a treatment that raises the function level from near dead to a state of reduced activity might be more cost effective than a treatment that raises someone from reduced activity to a higher level of functional ability; it is the degree of improvement rather than the level of function that is the underlying preference, an ethical assumption not yet examined.

Cost studies also assume that individual need preferences are fairly expressed through aggregate measure. This also unexamined assumption takes the position that individual variation in preference should not take precedence over averages of large groups. In developing health policy, this assumption is unavoidable. Nevertheless, society clearly indicates that on occasion it is willing to make exceptions. For example, a young girl trapped in a well demands the country's attention and millions of dollars worth of medical care, whereas the same girl could have difficulty obtaining prenatal care, vaccinations, and varied health benefits that would be measured in dollars. Similarly, whereas there is current debate over organ transplantation, occasional liver transplants have been carried out because of media and presidential attention. Of these exceptional circumstances, society may be making a sentimental value choice that, on occasion, if an individual evokes or appeals to an ideal (eg, human kindness, charity), society is willing to put aside rational calculations. Thus, society demonstrates that it is inconsistent in establishing ethical norms with respect to the value of human life—an inconsistency that is the nemesis of cost-effectiveness research.

CONCLUSIONS

Cost studies are not panaceas; they have definite limitations. Some of these limitations result from controversies surrounding the practical, theoretical, and ethical issues described in this paper. While some of the ethical consid-

erations can be addressed by modifying the approach to the analysis, others mark the center of important controversies in the field. More work is needed in these and other issues central to the current concepts of cost effectiveness.

Many elements of cost-effectiveness research are not controversial. As techniques improve and cost studies become more popular, they are relied upon to a greater degree when deciding important issues of resource allocation and so-called proper medical practice. A cost study improperly performed, or one used to make policy statements beyond those appropriate considering the limitations of the study, may lead to incorrect, potentially hazardous policy statements. It is therefore essential that all cost studies adhere to certain basic principles: clarifying the perspective of the study; including the direct costs of the program, the costs of future workup or treatment, the costs of evaluating and treating complications resulting from the program, and the future medical costs averted; discounting future costs and health effects; and performing a sensitivity analysis.

References

1. Walgren D: Medicare coverage for cyclosporine. Med Decis Making 1987; 7:1–2
2. Doubilet P, Weinstein M, McNeil B: Use and misuse of the term cost effective in medicine. N Engl J Med 1986; 314:253–256
3. Weinstein MC, Fineberg HV (eds): Clinical Decision Analysis. Philadelphia, WB Saunders, 1980; pp 254–263, 300–302
4. Russell LB: Evaluating Preventive Care. Washington, DC, The Brookings Institution, 1987, pp 70, 73
5. Chen M, Bush J, Patrick D: Social indicators for health planning and policy analysis. Policy Sci 1975; 6:71–89
6. Kaplan R, Bush J: Health-related quality of life measurement for evaluation research and policy analysis. Health Psychol 1982; 1: 61–80
7. Bush JW: General health policy model/quality of well-being (QWB) scale. In Wenger NK, Mattson ME, Furberg CD, Elinson J (eds): Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies. New York, Le Jacq Publishing, 1984, pp 189–199
8. Weinstein M, Stason WB: Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med 1977; 296:716–721
9. World Health Organization. Health Promotion: A Discussion Document on the Concept and Principles. Copenhagen, WHO Regional Office for Europe, 1984
10. Collen MF: The baseline screening electrocardiogram: Is it worthwhile? An affirmative view. J Fam Pract 1987; 25:393–396
11. Dervin JV: Feasibility of 105-cm flexible sigmoidoscopy in family practice. J Fam Pract 1986; 23:341–344
12. Estes, EH Jr: The baseline screening electrocardiogram: Is it worthwhile? An opposing view. J Fam Pract 1987; 25:395–396
13. Kaplan RM, Bush JW, Berry CC: Health status: Types of validity and the index of well-being. Health Serv Res 1976; 11:478–507
14. Bush JW, Fanshel S, Chen M: Analysis of a tuberculin testing program using a health status index. J Socio-Econ Plann Sci 1972; 6:49–69
15. Bush JW, Chen M, Patrick DL: Cost-effectiveness using a health status index: Analysis of the New York State PKU screening program. In Berg R (ed): Health Status Indexes. Chicago, Hospital Research and Educational Trust, 1973, pp 172–208
16. Epstein KA, Schneiderman LJ, Bush JW, Zettner A: The 'abnormal'; screening serum thyroxine (T4): Analysis of physician response, outcome, cost and health effectiveness. J Chronic Dis 1981; 34:175–190
17. Anderson JP, Moser RJ: Parasite screening and treatment among Indochinese refugees: Cost – benefit/utility and the General Health Policy Model. JAMA 1985; 253:2229–2235
18. Weinstein MC: Challenges for cost-effectiveness research. Med Decis Making 1986; 6:194–198
19. Weinstein MC: Estrogen use in post-menopausal women—Costs, risks and benefits. N Engl J Med 1980; 303:308–316
20. Chen M, Bush JW: Maximizing health system output with political and administrative constraints using mathematical programming. Inquiry 1976; 13:215–227
21. Chen M, Bush JW, Patrick DL: Social indicators for health planning and policy analysis. Policy Sci 1975; 6:71–89
22. Bush JW, Blischke WR, Berry CC: Health indices, outcomes, and quality of care. In Yaffe R, Zalkind D (eds): Evaluation of Health Services Delivery. New York, Engineering Foundation, 1975, pp 313–339
23. Anderson J, Bush J, Chen M, Dolenc D: Policy space areas and properties of benefit – cost/utility analysis. JAMA 1986; 255: 794–795
24. Sackett DL, Haynes RB, Tugwell P (eds): Clinical Epidemiology: A Basic Science for Clinical Medicine. Boston/Toronto, Little, Brown, 1985, p 314
25. Fuchs VR, Zeckhauser R: Valuing health—A 'priceless' commodity. AEA Pap Proc 1987; 77:263–268
26. Keeler EB, Cretin S: Discounting of life-saving and other nonmonetary effects. Manage Sci 1983; 29:300–306
27. Sacks HS, Berrier J, Reitman D, et al: Meta-analysis of randomized controlled trials. N Engl J Med 1987; 316:450–455
28. Weinstein MC: Economic assessments of medical practices and technologies. Med Decis Making 1981; 1:309–330
29. Cocchi SL, Broome CV, Hightower AW: Immunization of US children with Hemophilus influenzae type b polysaccharide vaccine. JAMA 1985; 253:521–529
30. Wulff HR: How to make the best decision: Philosophical aspects of clinical decision theory. Med Decis Making 1981; 1:277–283
31. Rawls J: A Theory of Justice. Cambridge, Harvard University Press, 1971