

How to read clinical journals: VII. To understand an economic evaluation (part B)*

DEPARTMENT OF CLINICAL EPIDEMIOLOGY AND BIOSTATISTICS, MCMASTER UNIVERSITY HEALTH SCIENCES CENTRE

The readers of clinical journals increasingly encounter articles on the economic evaluation of one or more clinical maneuvers or programs and are often faced with the task of assessing their results. As shown in the case presentations in part A, the question that readers of such articles are most likely to ask themselves is Are these results useful to me in my setting? The answer is determined by the answers to the following specific questions:

- Are the methods employed in the study appropriate, and are the results valid?

- If the results are valid, would they apply to my setting?

This second in the pair of articles on understanding economic evaluation relates mainly to the former question. It is designed to assist users of economic evaluations in assessing the validity of the results they encounter.

When assessing the validity of evidence, whether in terms of efficacy, effectiveness, availability or efficiency, we usually proceed by close-

ly examining the methods used to produce the evidence. Often it is helpful to separate the various elements of a method so that each can be scrutinized, a strategy we repeatedly applied in the previous clinical epidemiology rounds. Accordingly, we shall identify the key elements of an economic evaluation and, through the use of a set of readers' guides (Table I), discuss the methodologic standards that readers can expect to find in a well executed economic analysis. Of course, it is unrealistic to expect every study to satisfy all of the guides; however, the systematic application of the guides will allow readers to identify and assess the strengths and weaknesses of individual clinical studies.

Elements of a sound economic evaluation

1. Was a well defined question posed in answerable form?

Such a question will clearly identify the alternatives being compared and the viewpoint(s) from which the comparison is to be made. Questions such as Is a chronic home care program worth it? and Will a community hypertension screening program do any good? beg the further questions To whom? and Compared with what? Similarly, questions such as How much does it cost to run our intensive care unit? and What are the costs and outcomes of adolescent counselling by social workers? are not efficiency questions because they fail to specify the alternatives for comparison. (See part A for a review on the nature of economic evaluation.) This is not to say that

the answers to such questions will not provide important information on accounting or management. They may, but alone they do not qualify as efficiency statements.

A well specified question might be the following: From the viewpoint of (a) the budgets of both the Ministry of Health and the Ministry of Community and Social Services and (b) patients incurring out-of-pocket costs, is a chronic home care program preferable to the existing program of institutionalized extended care in designated wards of general hospitals? Note that the viewpoint for an analysis may be that of a specific provider or providing institution, a patient or groups of patients, a third-party (public or private) payer, or society (i.e., all costs and consequences to whomsoever they accrue). It may be that a program is preferable from the viewpoint of society but not from that of the providing institution. In such a case the Ministry of Health may wish to consider giving an incentive to the providing institution to ensure that the socially preferred program goes ahead. The existence of different viewpoints was highlighted by Weisbrod and colleagues¹ in their study of community-oriented and hospital-based treatments for patients with mental illness.

2. Was a comprehensive description of the competing alternatives given?

A clear and specific statement of the primary objective of each alternative program, treatment or service is critical in selecting the type of evaluation — cost-effectiveness,

*Parts I through V of this series were published in consecutive issues of the Journal starting with the Jan. 1, 1981 issue, and part VI was published in the Feb. 15, 1984 issue. Part VII(A) appeared in the June 1, 1984 issue.

Reprint requests to: Professor G.L. Stoddart or Professor M.F. Drummond, McMaster University Health Sciences Centre, Rm. 2C13, 1200 Main St. W., Hamilton, Ont. L8N 3Z5

Table 1—Detailed readers' guides for efficiency studies

1. Was a well defined question posed in answerable form?
 - (a) Did the study examine both costs and effects of the service(s) or program(s)?
 - (b) Did the study involve a comparison of alternatives?
 - (c) Was a viewpoint for the analysis stated or was the study placed in a particular decision-making context?
2. Was a comprehensive description of the competing alternatives given (i.e., can you tell who did what to whom where and how often)?
 - (a) Were any important alternatives omitted?
 - (b) Was (should) a "do-nothing" alternative (have been) considered?
3. Was there evidence that the programs' effectiveness had been established? Was this done through a randomized, controlled clinical trial? If not, how strong was the evidence of effectiveness?
4. Were all important and relevant costs and consequences for each alternative identified?
 - (a) Was the range wide enough for the research question at hand?
 - (b) Did it cover all relevant viewpoints (e.g., those of the community or society, patients and third-party payers)?
 - (c) Were capital costs as well as operating costs included?
5. Were costs and consequences measured accurately in appropriate physical units (e.g., hours of nursing time, number of physician visits, days lost from work or years of life gained) prior to valuation?
 - (a) Were any identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?
 - (b) Were there any special circumstances (e.g., joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?
6. Were costs and consequences valued credibly?
 - (a) Were the sources of all values (e.g., market values, patient or client preferences and views, policymakers' views and health care professionals' judgements) clearly identified?
 - (b) Were market values used for changes involving resources gained or used?
 - (c) When market values were absent (e.g., when volunteers were used) or did not reflect actual values (e.g., clinic space was donated at a reduced rate) were adjustments made to approximate market values?
 - (d) Was the valuation of consequences appropriate for the question posed (i.e., was the appropriate type, or types, of analysis — cost-effectiveness, cost-benefit or cost-utility — selected)?
7. Were costs and consequences adjusted for differential timing?
 - (a) Were costs and consequences that occurred in the future "discounted" to their present values?
 - (b) Was any justification given for the discount rate used?
8. Was an incremental analysis of costs and consequences of alternatives performed?

Were the additional (incremental) costs generated by the use of one alternative over another compared with the additional effects, benefits or utilities generated?
9. Was a sensitivity analysis performed?
 - (a) Was justification provided for the ranges of values (for key parameters) used in the sensitivity analysis?
 - (b) Were the study results sensitive to changes in the values (within the assumed range)?
10. Did the presentation and discussion of the results of the study include all issues of concern to users?
 - (a) Were the conclusions of the analysis based on some overall index or ratio of costs to consequences (e.g., cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?
 - (b) Were the results compared with those of other studies that had investigated the same questions?
 - (c) Did the study discuss the generalizability of the results to other settings and patient/client groups?
 - (d) Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g., distribution of costs and consequences or relevant ethical issues)?
 - (e) Did the study discuss issues of implementation, such as the feasibility of adopting the "preferred" program, given existing financial or other constraints, and whether any freed resources could be used for other worthwhile programs?

cost-benefit or cost-utility — to be undertaken. A full description of the alternatives is essential for three other reasons: (a) readers must be able to judge the applicability of the programs to their own settings, (b) readers should be able to assess whether any costs or consequences may have been omitted in the analysis, and (c) readers may wish to replicate the program procedures being described. Therefore, readers should be provided with information that allows identification of both the costs (Who does what to whom where and how often?) and the consequences (With what results?).

3. Was there evidence that the programs' effectiveness had been established?

We are not interested in the efficient provision of ineffective services (i.e., those that have been shown to do no more good than harm by themselves or compared with no treatment). In fact, we are not interested in the provision of such services under any conditions, efficient or otherwise. If something is not worth doing it's not worth doing well. Therefore, if the economic evaluation assumes effectiveness, some indication should be given of the prior validation of effectiveness. It is also possible that the efficiency evaluation may have been conducted simultaneously with the evaluation of efficacy or effectiveness. This is the case in many randomized trials of therapies that also include a comparison of the costs of the experimental program and a control, which may be a placebo or a currently existing program. Note, however, that efficiency evaluations alone cannot establish effectiveness. There are, after all, efficient methods of worsening the quality of life as well as improving it. (If you want to know more about the methods of determining whether a therapy does more good than harm you should read part V of this series.²)

4. Were all important and relevant costs and consequences for each alternative identified?

Even though it may not be possible or necessary to measure and value all of the costs and conse-

quences of the alternatives being compared, the important and relevant ones should be identified. The information in the viewpoint statement and program description should allow you to judge what specific costs and consequences or outcomes it is appropriate to include in the analysis.

An overview of the types of costs and consequences that may be relevant to economic evaluation of health services and programs is provided in Fig. 1. Three categories of costs are shown. Since the costs of a health care service or program are best thought of in terms of the resources used, category I contains the costs of organizing and operating the program. The identification of these costs often amounts to listing the "ingredients" of the program — both variable costs (such as those of health care professionals' time or supplies) and fixed or overhead costs (such as capital costs, rent and the costs of light and heat). These costs are often referred to by economists as "direct costs".*

Category II contains costs that are borne by patients and their fam-

ilies. These include any out-of-pocket expenses as well as the value of any resources they contribute to the treatment process. Patients and their families sometimes lose time from work while seeking treatment or participating in a health program. Such "production losses" are also a cost of the health care service or program and are often referred to by economists as "indirect costs". However, care must be taken when including this cost in an analysis, since its inclusion implies that the cost was incurred as a result of participation in treatment and therefore that the individual's condition alone would not have prevented productive activity.† Finally, the anxiety, and perhaps pain, associated with treatment constitutes a psychic cost frequently encountered by patients and their families.

While these two categories cover most of the costs relevant to economic evaluations of health care services, a third category also warrants mention. It may be that the operation of a health care service or program changes the use of resources in the broader economy outside the health sector. For example, an occupational health or safety pro-

gram may result in more costly production processes; thereby raising the price of, say, cars. In principle such instances should be identified, though in practice they may rarely be significant. (Few economic analyses of alternative health programs take them into account.)

Fig. 1 also shows three categories of consequences of health care services and programs. Category I contains therapeutic outcomes or effects of the alternatives. These effects will usually include changes in the physical, social or emotional functioning of individuals. In principle such changes can be measured objectively and refer only to an individual's ability to function and not to the significance, preference or value attached to this ability by the individual or by others.

The therapeutic effects of a service or program give rise to two other important categories of consequences. First, the effects may result in changes in the use of resources in the future (category II). Within the health care sector, less use of resources may be required for treatment of the condition and its sequelae than would otherwise have been the case. For example, an effective hypertension screening program averts the future cost of caring for stroke victims. The saving in the use of health care resources attributable to the screening program is usually referred to by economists as the

*Health care administrators sometimes reserve the term direct costs for variable costs only and may refer to overhead costs as indirect costs. In economic evaluations, however, economists use the term indirect costs to denote a separate and distinct type of cost, as we will explain later. Users of evaluations should be aware of this potential source of confusion.

†The complexity of the relation between lost work time and the value of forgone output places it beyond the scope and purpose of this article. However, for a discussion of its implications for categories II and III you can read Stoddart's article.³

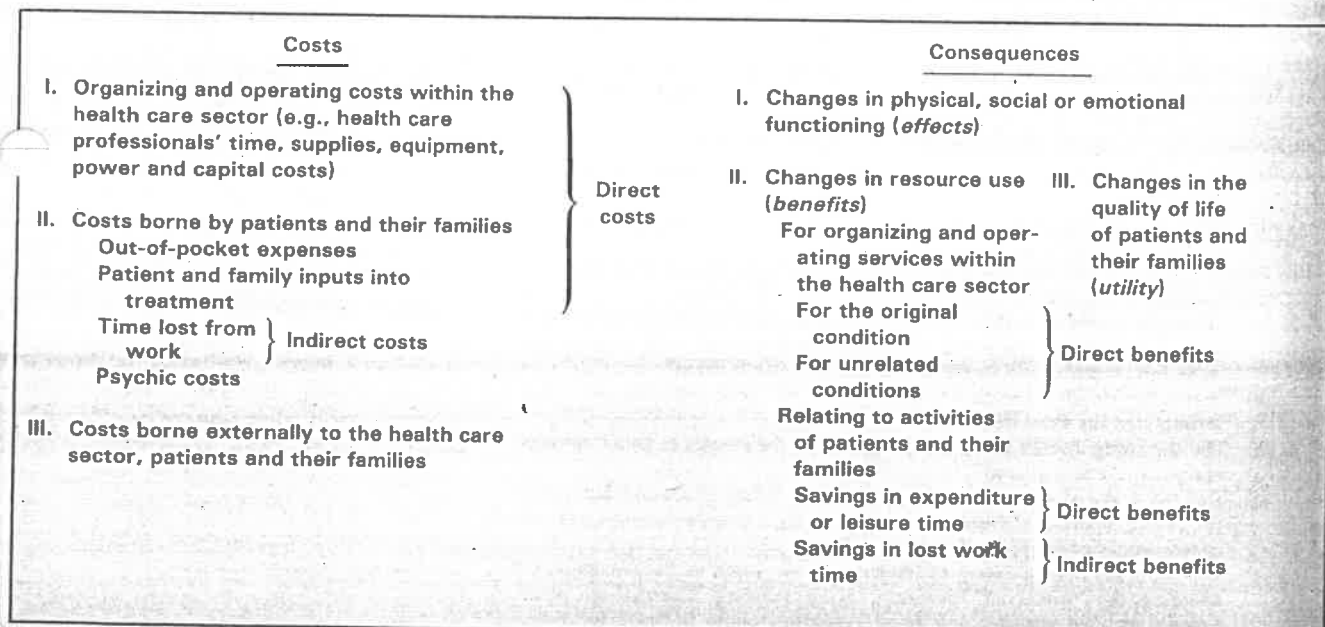


Fig. 1—Types of costs and consequences relevant to economic evaluation of health care services and programs (adapted, with permission, from reference 3).

direct benefit of the screening program. Notice, however, that if we adopt the viewpoint of a health care system the direct benefits are sometimes negative owing to the increased use of services for the treatment of conditions (e.g., arthritis) that may develop in patients during their newly prolonged lives. The therapeutic effects of a health care service or program may also affect the use of resources by patients and their families. Of particular interest is the possibility that patients and their families may gain working time as a result of their participation in treatment. These production gains are usually referred to by economists as "indirect benefits".

The inclusion of indirect benefits in economic evaluations is a source of some controversy among analysts. It is sometimes argued (rather narrowly it seems) that health care evaluation should confine itself to changes in the use of resources in the health care sector only, rather than in the entire economy. More serious is the assertion that changes in the output of individuals or groups are simply not the grounds upon which we usually make decisions about allocation of health care resources that will affect those individuals or groups. Therefore, it is misleading to enter the value of such changes into a cost-benefit calculation. A third criticism is that the valuation of indirect benefits (usually through increased earnings of individuals) makes a series of value judgements and assumptions that may only be appropriate in a limited number of cases. While it is not possible to discuss and evaluate these claims here, you should be aware that the inclusion of indirect benefits in a cost-benefit analysis may not be straightforward.^{4,5*}

The therapeutic effects of health care services and programs also give rise to another extremely important category of consequences; namely, changes in the quality of life of patients and their families (category

III). The change in quality of life produced by the therapeutic effects is distinguished from the effects themselves by the significance or value that patients and their families attach to the effects. It is, of course, possible — and, in fact, likely — that different individuals place a different importance on the same level of physical, social or emotional functioning. (This was demonstrated by the example of the impact of a broken arm on the sign painter and the translator in part A.)

With respect to both the costs and the consequences we have described it may be unrealistic to expect all relevant items to be measured and valued in an economic analysis owing to the small impact of some relative to the effort required to measure or value them accurately; however, it is helpful to users to have as many of them as possible identified. It is particularly important that the outcomes of interest be identified clearly enough for you to judge the appropriateness of the type(s) of economic evaluation chosen. That is, it should be apparent (a) whether a single outcome is of primary interest as opposed to a set of outcomes that are each of some, if not equal, interest, (b) whether the outcomes are common to the two alternatives being compared, and (c) to what degree each program is successful in achieving each outcome of interest. Similarly, it is important to know whether the consequence of primary interest is the therapeutic effect (which implies that a cost-effectiveness analysis should be done if possible), the net change in the use of resources (cost-benefit analysis) or the quality of life of the patients and their families (cost-utility analysis).

5. *Were costs and consequences measured accurately in appropriate physical units?*

While identification, measurement and valuation often occur si-

multaneously in economic analyses, you should view each as a separate phase of the analysis. Once the important and relevant costs and consequences have been identified, they must be measured in appropriate physical or natural units. For example, measurement of the operating costs of a particular screening program may yield a partial list of "ingredients" as follows: 500 physical examinations performed by physicians, 10 weeks of salaried nursing time, 10 weeks' use of a 100-m² clinic, 20 hours of medical librarian research time from an adjoining hospital and so forth. Similarly, costs borne by patients may be measured by the amount of medication purchased, by the number of times travel was required for treatment or by the time lost from work while the patient was being treated.

Situations in which resources are used jointly by one or more programs present a particular challenge to accurate measurement. How much use of a resource should be allocated to each program and on what basis? A common example is found in every hospital, where numerous clinical services and programs share overhead services (e.g., power, cleaning and administration) that are provided centrally. In general, there is no nonarbitrary solution to the measurement problem. You should, however, satisfy yourselves that "reasonable" criteria (e.g., number of square metres, number of employees or number of cases) have been used in the distribution of the common costs. You should definitely ascertain that such shared costs have, in fact, been allocated to participating services or programs, as this is a common omission in evaluations. Clinical service directors often argue that small changes in the size of their programs (up or down) do not affect the use of central services. Sometimes it is even argued that overhead costs are unaffected by the service itself. However, though this argument may be appealing from the viewpoint of a particular program or service director, the extension of this method to each service in the hospital would imply that all the services could be operated without light, heat, power and secretaries!

services undoubtedly exist, by their very nature they are difficult to include in a cost-benefit analysis, which expresses costs and consequences in dollars. They presumably are taken into account, however, in cost-utility analyses, wherein program effects are translated into a measure of value based on preferences rather than dollars, as discussed in question 6.

*Those who criticize the inclusion of indirect benefits, saying You can value a livelihood but you can never value a life! appear to be confusing indirect benefits with another type of benefit. This is the intangible value we, as individuals and as a society, place on life itself (regardless of earning potential) and on the avoidance of pain and suffering. Although intangible benefits and costs of health care

With respect to the measurement of consequences, if the outcomes of interest have been clearly identified, then the selection of appropriate units of measurement for program effects should be relatively straightforward. Effects might relate to mortality and be measured in years of life gained or deaths averted, or they might relate to morbidity and be measured in reductions in the number of days of disability or improvements in health status according to some index of physical, social or emotional function. They may be even more specific, depending on the alternatives under consideration. Thus, "percentage increase in weight-bearing ability" may be an appropriate natural measurement unit for an evaluation of a physiotherapy program, and "the number of correctly diagnosed cases" may be appropriate for a comparison of enography with leg scanning in the diagnosis of deep-vein thrombosis.

Changes in the use of resources resulting from the effects will be measured in physical units similar to those used for costs. Thus, the changes in use resulting from any particular program will likely be recorded in numbers of procedures or in amounts of time, space or equipment. Changes in the use of resources by patients will continue to be measured, for example, in amounts of medication purchased or number of trips taken for treatment.

While the nature of changes in the quality of life may be described in an economic evaluation, measurement in objective, physical or natural units is difficult, although the consequence of some surgical interventions may be quantified in "number of complications". However, the adjustment of effects for quality of life is usually a matter of valuation.

6. Were costs and consequences valued credibly?

The sources and methods of valuation of costs, benefits and utilities should be clearly stated in an economic evaluation. Costs are usually valued in units of local currency on the basis of prevailing "prices" of personnel, commodities, services and so forth and can often be taken directly from program budgets. All current and future program costs

are usually valued in constant dollars of a base year (usually the present) to remove the effects of inflation from the analysis.

It should be remembered that the object in valuing costs is to obtain an estimate of the worth of the resources used by the program. This may necessitate adjustments to some apparent program costs, as in the case of volunteer labour or subsidized services received by one program instead of another. In addition, valuation of the cost of a day of institutional care for a specific condition is particularly troublesome, in that the use of an average cost per day (the widely quoted "per diem"), calculated on the basis of the institution's annual caseload, almost certainly overestimates or underestimates (sometimes by a large amount) the actual cost for any specific condition. You should thus approach per-diem values with extreme caution.*

Valuation of direct benefits proceeds in the same fashion as that of costs and is subject to the same caveats since the benefits are usually the expected future costs that are saved. Valuation of production gains or indirect benefits (i.e., changes in the value of output of individuals or groups who receive the health care program or service) usually employs the wage rates for individuals or groups to value the increased working time available. It is here that critics of cost-benefit analyses point out the inequity associated with

*In principle, and with great effort in practice, it is possible to identify, measure and value each resource (e.g., drugs, nursing time, light or food) used in treating a specific patient or group of patients. While this yields a relatively accurate cost estimate, the detailed monitoring and data collection are usually prohibitively expensive. The other broad costing strategy is to start with the institution's total costs for a particular period and improve upon the method of simply dividing by the total number of patient days to produce an average cost per day. Quite sophisticated methods of cost allocation to individual hospital departments or wards have been developed.⁶ An intermediate method involves acceptance of the components of the general per diem relating to "hotel" costs (since these are relatively invariant for all patients) combined with more precise calculation of the costs of medical treatment of the individual patients.⁷ Of course, the effort devoted to estimating accurate per diems depends on their overall importance in the study; however, unthinking use of per diems or average costs should be avoided.

linking estimates of the value of health care programs to the vagaries of the market. They argue that acceptance of existing wage rates coupled with the inclusion of indirect benefits biases cost-benefit studies against programs aimed at minority groups, housewives, the elderly, children and the unemployed. Although it may be possible to adjust some of the estimates to acknowledge this problem (e.g., by imputing a value to housewives' services based on wages for similar work) the indirect benefit issue remains controversial.

In valuation of preferences or utilities we are basically attempting to ascertain how much better, all things considered, the quality of life is in one health situation or state (e.g., when dialysis is performed at home with help from a spouse or a friend) compared with another (e.g., when dialysis is performed in hospital). Several techniques are available for making the comparison; each will produce a utility value (mentioned in part A) with which one can increase or decrease the value of time spent in health situations resulting from the alternative in question relative to a baseline. Usually the results of utility analyses will be expressed in "healthy days" or "quality-adjusted life-years" resulting from the programs being evaluated.

Two broad approaches to utility analysis can be found in the literature. The first approach, outlined by Torrance,⁸ emphasizes the development of *measurement* methods and empirical testing in different populations. The other approach, outlined by Weinstein,⁹ emphasizes the estimation of utility values by a quick (and inexpensive) *consensus-forming* exercise and then the performance of an extensive sensitivity analysis on the chosen values to see whether the results change if the utility estimates are varied. We see a role for both approaches. The latter is useful in persuading decision-makers to think about problems in allocating resources and is, in fact, relatively quick and inexpensive. The measurement approach is useful in highlighting the fact that different individuals (doctors, policymakers, patients and the general public as taxpayers) may have dif-

ferent values, and it is clearly crucial in situations in which the result is sensitive to the utility values assigned. (An example of such a situation arose in the study by Stason and Weinstein¹⁰ on the economics of hypertension therapy. The result of their study was sensitive to whether it was assumed that the side effects of antihypertensive drugs constituted a 1% or a 2% reduction in health status.)

Since the measurement of preferences in health is relatively new, there are naturally many unresolved issues in cost-utility analyses. You will probably want to know, at the least, *whose* preferences — patients', providers', taxpayers' or bureaucrats' — were used to construct the utility values. If patients' preferences were not used you may want further assurance that the persons whose preferences were used clearly understood the characteristics of the health state, either through personal experience or through a description of the state presented to them.

7. Were costs and consequences adjusted for differential timing?

Since a comparison of programs or services must be made at one point in time (usually the present), the timing of program costs and consequences that do not occur entirely in the present must be taken into account. Different programs may have different time profiles of costs or consequences. For example, the primary benefits of an influenza immunization program are immediate, whereas those of a hypertension screening program will occur in the future. The time profile of costs and consequences may also differ within a single program; the costs of the hypertension screening program would be incurred in the present. Future dollar costs and benefits are therefore reduced or "discounted" to reflect the fact that dollars spent or saved in the future should not weigh as heavily in program decisions as dollars spent or saved today. This is primarily due to the existence of "time preference" — that is, we, as individuals and as a society, prefer to have dollars or resources now as opposed to later because we can use them in the interim. This is evi-

denced by the existence of interest rates (as well as by the popular wisdom about "a bird in the hand"). Moreover, since time preference is not exclusively a financial concept, outcomes should also be discounted in cost-effectiveness and cost-utility studies.¹¹⁻¹⁴

8. Was an incremental analysis of costs and consequences of alternatives performed?

For a meaningful comparison it is necessary to examine the additional costs imposed by the use of one service or program over another, compared with the additional effects, benefits or utilities it delivers. This "incremental" approach to the analysis of costs and consequences can be illustrated by one of the examples cited in part A of this article; namely, the strategies for diagnosing deep-vein thrombosis.⁷ Table II shows the costs and outcomes (in terms of correct diagnoses*) generated by two alternative strategies: impedance plethysmography alone and impedance plethysmography plus outpatient venography if the former gives negative results.⁷ Impedance plethysmography is a noninvasive strategy, whereas venography, the diagnostic "gold standard" for deep-vein thrombosis, can cause pain and other side effects. Although one could compare the simple ratios of costs to out-

*The study by Hull and associates⁷ is an example of a cost-effectiveness analysis in which the outcomes are not therapeutic effects but, rather, intermediate diagnostic outcomes with direct implications for therapeutic effects in that the failure to diagnose deep-vein thrombosis leads directly to increased morbidity and mortality.

comes for the two alternatives one should compare the *incremental* costs with the *incremental* outcomes, since this will tell us how much we are paying for each extra correct diagnosis in adding the extra diagnostic test. In this case the relevant figure is \$4781 per correct diagnosis, not the average figure for the second program, which is \$3003 per correct diagnosis. It may be decided that \$4781 is still a price worth paying; however, it is important to be clear on the principle since, as we pointed out in part A, in screening for cancer of the colon there was a big difference between the average cost per case detected of a protocol of six sequential tests and the incremental cost of performing a sixth test, having already done five.¹⁵

Obviously similar analyses could be performed if the consequences were effects in natural units (e.g., years of life) or in utilities (e.g., quality-adjusted life-years).

9. Was a sensitivity analysis performed?

Every evaluation will contain some degree of uncertainty, imprecision or methodologic controversy. What if the compliance rate for influenza vaccination was 10% higher than that considered for the analysis? What if the hospital per diem still understated the true resource cost of a treatment program by \$100? What if a discount rate of 6% instead of 2% had been used? What if indirect costs and benefits had been excluded from the analysis? You will often ask these and similar questions; therefore, careful analysts will identify critical methodologic

Table II—Economic evaluation of alternative diagnostic strategies for 516 patients with clinically suspected deep-vein thrombosis⁷

Program*	Cost (\$US)	Outcome (no. of correct diagnoses)	Ratio of cost to outcome (cost [\$US] per correct diagnosis)
IPG alone	321 488	142	2264
IPG and outpatient venography if results of IPG are negative	603 552	201	3003
Increment of second program over first program	282 064	59	4781

*IPG = impedance plethysmography.

assumptions or areas of uncertainty. Furthermore, they will often attempt to rework the analysis (qualitatively at least, if not quantitatively) with different assumptions or estimates to test the sensitivity of the results and the conclusions to such changes. If large variations in the assumptions or estimates underlying an analysis do not produce significant alterations in the results, then one would tend to have more confidence in the original results. If the converse occurs, then more effort is required to reduce the uncertainty or improve the accuracy of the critical variables. In either case a sensitivity analysis is an important element of a sound economic evaluation.

10. Did the presentation and discussion of the results of the study include all issues of concern to users?

It will be clear from the preceding discussion that economic analysts have to make many methodologic judgements when undertaking a study. Faced with users who may be mainly interested in the "bottom line" (e.g., should they buy a computerized tomography scanner?) how should they present their results?

Decision indices, such as cost-effectiveness and cost-benefit ratios, are useful in summarizing the results of a study. However, they should be used with care, for the reader, when interpreting them, may not be completely clear on what has gone into their construction. Some analysts give a range of results. For example, in an economic evaluation of neonatal intensive care for very-low-birth-weight infants Boyle and coworkers⁶ compared the results for infants weighing less than 1000 g and from 1000 to 1500 g in terms of costs up to the time of hospital discharge, costs and consequences to age 15 years and costs and consequences for a lifetime (Table III). They left it to the reader to decide which index (or indices) to use in judging neonatal intensive care, since the different measures incorporate different value judgements and amounts of precision. (For example, the index of net economic benefit includes production gains or

losses, and the index of cost per quality-adjusted life-year incorporates the preferences for health states of a sample of the local population.)

This leads to another general point; namely, it is important for analysts to be as explicit as possible about the various judgements they have made in carrying out the study. A good study should leave you more (rather than less) aware of the various technical and value judgements necessary to arrive at decisions on the allocation of resources in health care.

Finally, a good study should begin to help you interpret the results in the context of your own situation. This can be done if the analysts are explicit about the viewpoint for the analysis and indicate how particular costs and benefits might vary by location. For example, the costs of instituting day-care surgery may vary depending on whether a purpose-built day-care unit already exists or whether wards have to be converted. Similarly, the benefits of day-care surgery may vary depending on whether there is a shortage of beds and on whether the beds will be left empty.¹⁶ Obviously it is impossible for the analysts to anticipate every possibility in every location, but one limitation of economic evaluation techniques is that they assume that freed resources will be put to other beneficial uses. Evans and Robinson¹⁷ argue that in the case of

day-care surgery the full economic payoff may not have been obtained in at least one Canadian hospital.

Limitations of economic evaluation techniques

The main purpose of this pair of articles is to make you more aware of the methodologic judgements involved in an economic evaluation in the health care field. In Table I we have consolidated the points made in the text into a checklist of questions you should ask when critically assessing the results of an economic evaluation. Some of the questions signal limitations of economic evaluation techniques. For example, economic evaluation techniques assume, rather than establish, program effectiveness. There are several other limitations of which you should be aware.

Of primary concern from a policy viewpoint is that economic evaluations do not usually incorporate into the analysis the importance of the distribution of costs and consequences. Yet in some cases the identity of the recipient group (e.g., the poor, the elderly or working mothers) may be an important factor in assessing the social desirability of a service or program. Indeed, it may be the motivation for the program. Although it is sometimes suggested that differential weights be attached to the value of outcomes to special

Table III—Measures of economic evaluation of neonatal intensive care, according to birth-weight class (5% discount rate)*

Period	Birth-weight class; cost (\$)†	
	1000–1499 g	500–999 g
To hospital discharge‡		
Cost/additional survivor at hospital discharge	59 500	102 500
To age 15 years (projected)		
Cost/life-year gained	6 100	12 200
Cost/QALY§ gained	7 700	40 100
To death (projected)		
Cost/life-year gained	2 900	9 300
Cost/QALY gained	3 200	22 400
Net economic benefit (loss)/live birth	(2 600)	(16 100)
Net economic cost/life-year gained	900	7 300
Net economic cost/QALY gained	1 000	17 500

*Adapted, with permission, from reference 6.

†In 1978 Canadian dollars; multiply by 0.877 to calculate equivalent cost in 1978 US dollars.

‡All costs and effects occurred in year 1.

§QALY = quality-adjusted life-year.

recipient groups, this is not usually done in an economic evaluation. Rather, an "equitable" distribution of costs and consequences across socioeconomic or other defined groups is viewed as a competing dimension, in addition to efficient deployment of resources, upon which decisions are made.

As we have already pointed out, economic evaluation techniques assume that resources that are freed up or saved by preferred programs will not, in fact, be wasted but will be used in alternative worthwhile programs. This assumption warrants careful scrutiny, for if the freed resources are consumed by ineffective or unevaluated programs, then not only is there no saving, but the overall costs of the health care system will actually increase without any assurance of additional improvements in the health status of the population.

Finally, an evaluation of any sort is costly. If we bear in mind that even a cost-benefit analysis should be subject to a cost-benefit analysis, it seems reasonable to suggest that economic evaluation techniques will prove most useful when program objectives require clarification, when the competing alternatives are significantly different in nature and when large resource commitments are being considered.

Conclusion

In this pair of articles we have

tried to help potential users of economic evaluations to understand such studies and assess their validity and usefulness. Our intent has not been to create hypercritical users who will be satisfied only by superlative studies. As we have emphasized, it is unlikely that every study will satisfy all the methodologic criteria we have discussed. However, the readers' guides should help you to quickly identify the strengths and weaknesses of any study.

We thank our colleagues and students for their suggestions and criticisms of earlier versions of this paper.

References


1. WEISBROD BA, TEST MA, STEIN LI: Alternative to mental hospital treatment. II. Economic cost-benefit analysis. *Arch Gen Psychiatry* 1980; 37: 400-405
2. Department of clinical epidemiology and biostatistics, McMaster University Health Sciences Centre: How to read clinical journals: V. To distinguish useful from useless or even harmful therapy. *Can Med Assoc J* 1981; 124: 1156-1162
3. STODDART GL: Economic evaluation methods and health policy. *Eval Health Prof* 1982; 5: 393-414
4. DRUMMOND MF: Welfare economics and cost-benefit analysis in health care. *Scott J Polit Econ* 1981; 28: 125-145
5. WILLIAMS AH: Welfare economics and health status measurement. In VAN DER GAAG J, PERLMAN M (eds): *Health, Economics and Health Economics*, Elsevier, Amsterdam, 1981: 271-281
6. BOYLE MH, TORRANCE GW, SINCLAIR JC, HORWOOD SP: Economic evaluation of neonatal intensive care of very-low-birth-weight infants. *N Engl J Med* 1983; 308: 1330-1337
7. HULL R, HIRSCH J, SACKETT DL, STODDART G: Cost effectiveness of clinical diagnosis, venography and noninvasive testing in patients with symptomatic deep-vein thrombosis. *N Engl J Med* 1981; 304: 1561-1567
8. TORRANCE GW: Preferences for health states: a review of measurement methods. In SINCLAIR JC (ed): *Clinical and Economic Evaluation of Perinatal Programs. Proceedings of Mead Johnson Symposium on Perinatal and Developmental Medicine, Number 20, Vail, Colorado, June 6-10, 1982*, Mead Johnson, Evansville, Ind, 1982: 37-45
9. WEINSTEIN MC: Economic assessments of medical practices and technologies. *Med Decis Making* 1981; 1: 309-330
10. STASON WB, WEINSTEIN MC: Allocation of resources to manage hypertension. *N Engl J Med* 1977; 296: 732-739
11. WEINSTEIN MC, STASON WB: Foundations of cost-effectiveness analysis for health and medical practices. *Ibid*: 716-721
12. DRUMMOND MF: *Principles of Economic Appraisal in Health Care*, Oxford U Pr, Oxford, 1980
13. Idem: *Studies in Economic Appraisal in Health Care*, Oxford U Pr, Oxford, 1981
14. WARNER KE, LUCE BR: *Cost-benefit and Cost-effectiveness Analysis in Health Care: Principles, Practice and Potential*, Health Admin Pr, Ann Arbor, Mich, 1982
15. NEUHAUSER D, LEWICKI AM: What do we gain from the sixth stool guaiac? *N Engl J Med* 1975; 293: 226-228
16. RUSSELL IT, DEVLIN HB, FELL M, GLASS NJ, NEWELL DJ: Day-case surgery for hernias and haemorrhoids: a clinical, social, and economic evaluation. *Lancet* 1977; 1: 844-847
17. EVANS RG, ROBINSON GC: Surgical day care: measurements of the economic pay-off. *Can Med Assoc J* 1980; 123: 873-880

Varia

IN ANGINA AND
HYPERTENSION.

Corgard (nadolol)
ONCE A DAY

because
renal blood flow
is important

 SQUIBB

Index Medicus

Index Medicus is the basic tool for medical literature searches. It is the National Library of

Medicine's major index to nearly 2 700 journal titles. It is published monthly and is cumulated yearly. Review articles are included in the subject section. It is subdivided into separate author and subject sections. Medical literature, beginning in 1879 with the old *Index Medicus*, is available through a series of indexes to the less current literature. The original *Index Medicus* was edited for many years by John Shaw Billings and Robert Fletcher. In 1960, the new, ongoing *Index Medicus* was begun.