Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem

Amiram Gafni, PhD; Stephen Birch, DPhil

Laupacis and associates[1] proposed tentative guidelines for using clinical and economic evaluations of health care services. The motivation for the development of such guidelines was as follows.

> For the results of clinical and economic evaluations to be used for policy formulation it is important to develop an idea of the orders of magnitude of cost-effectiveness that are likely to be associated with wise adoption and utilization and with unwise use of health care resources.

Five grades of recommendations (A to E) were proposed to classify technologies: a grade A technology is more effective and less costly than the existing technology, whereas a grade E one is less or equally effective but more costly. Grades B through D define technologies that are more effective but more costly, with an increasing incremental cost per quality-adjusted life-year (QALY).

In this paper we aim (a) to identify and challenge the main assumptions on which the guidelines were based and the policy implications of their adoption and (b) to describe an alternative approach for evaluating the appropriateness of the adoption of a new technology — a practical step toward improving the efficient use of health care resources.

Do the guidelines relate to the goals?

Laupacis and associates argued that their proposed guidelines on thresholds of cost–utility ratios allow scarce resources to be allocated for the maximum clinical benefit. They emphasized that the guidelines "do not directly address . . . how much in aggregate Canada should spend on health care. Their main purpose is to assist in deciding which technologies and programs should be funded *within any given budget* [emphasis added]."

This position is consistent with the traditional rationale for cost-effectiveness analysis: "For any level of resources available, society (or the decision-making jurisdiction involved) wishes to maximize the total aggregate health benefits conferred."[2]

Laupacis and associates, however, did not justify the use of their guidelines as a means of achieving or approaching their stated goal. They argued that there are "compelling reasons" for introducing a grade A technology (it is more effective and less costly than the current technology), but they did not consider the distributional and hence the social welfare consequences of this decision; they assumed that the adoption of a grade A technology implied no adverse redistribution of funds within society.

The "getting-more-for-less" situation is relatively rare in practice, because most new technologies tend to be more effective but more costly (grades B through D). One possible reason is that the providers of the "more effective" programs try to capture some of the public's value for the additional benefit by setting prices accordingly.

In this article, for simplicity, we focus on grade B technologies (those costing less than $20 000 per

Dr. Gafni is professor and Dr. Birch associate professor, Centre for Health Economics and Policy Analysis, Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ont.

Reprint requests to: Dr. Amiram Gafni, Centre for Health Economics and Policy Analysis, Department of Clinical Epidemiology and Biostatistics, McMaster University, 1200 Main St. W, Hamilton, ON L8N 3Z5

← For prescribing information see page 1061

QALY). According to the proposed guidelines "technologies that cost less than $20 000 per QALY are almost universally accepted as being appropriate ways of using society's and the health care system's resources."[1] These cost-utility thresholds were selected according to "previously suggested guidelines," which were based on arbitrary rather than objective criteria. A more important point is that the adoption of these guidelines will increase health care costs in producing the additional benefits. This is inconsistent with the authors' stated purpose of maximizing benefits by carefully selecting programs "within any given budget."

If a ministry of health followed the guidelines and adopted a grade B technology, the resources released by eliminating the existing program would not cover the costs of the new technology, and extra funds would need to be found. Laupacis and associates recognized this paradox and suggested that "if, for instance, most technologies were classified as grade A or B the implication might be that health care would warrant an increase in the aggregate level of expenditures." Although the guidelines were intended to assist with making decisions within a given budget, they appear to have been used to justify an increase in that budget. No consideration is given to the source of additional funds (e.g., other social programs, private consumption or borrowing against future wealth) or the forgone benefits of this required shift in resources. This is fundamental, however, to maximizing clinical benefits.

The proposed guidelines assume that a pool of underemployed or inefficiently employed resources, although not identified, exists somewhere within the current health care system and that this pool can be drawn upon to provide the additional resources. Later, Laupacis and associates introduced an additional consideration, the "tiebreaker": a situation in which a choice is to be made between two new, alternative technologies having *equal* cost-effectiveness ratios in relation to the total size of the programs, the preferred technology being the least costly. In other words, they recognized the opportunity cost implications of drawing resources from the unknown residual pool when more than one new program is evaluated. However, when one new technology is considered in isolation, such reservations seem to disappear (e.g., as in the claim that grade B technologies are almost universally accepted and warrant adoption and utilization).

Proponents of a particular program can enhance the apparent attractiveness (cost-effectiveness) of their program by creative choice of the comparator. For example, in the published "league tables" of cost-utility ratios, renal dialysis appears to be far less attractive than smoking cessation; however, the two programs are not compared with the same "existing" use of resources, and hence the comparison is not valid. Of course the same, fallacious, "economic" arguments could be used by those concerned with increasing expenditures for education, housing or private consumption, whereby a new program with an incremental cost per QALY ratio of less than an arbitrarily determined level should be adopted regardless of the identification of the required additional funds. We suspect that those actively involved in the health care system would not look upon such potential threats to the budget too kindly, and justifiably so.

Clearly, if economic principles are important in allocating funds within a given health care budget, then analysts might consider reviewing the economics literature for a discussion of these problems and methods for determining optimum allocation.[3,4]

Choice of clinical outcomes

Laupacis and associates argued that "in general, measurements of patient or societal preferences are preferred for assessing health care technologies and forming policy." QALYs are considered a utility-based index of preference that can be used as a "'common yardstick' . . . to compare the effectiveness of various interventions."[1] However, the authors recognized several limitations of QALY measures (as opposed to utility measures). In particular, the QALY assumes that the "value" attributed by a person to a particular health state is independent of the duration of that state, the state of health before the onset of the current state and health states expected in the future. It also assumes that individuals' preferences are such that the conversion of time in "ill health" to time in "full health" is linearly related to the time spent in ill health. Yet these and other assumptions were invalidated on the basis of empirical evidence[5] and can lead to biased estimates of the magnitude of a program's effects and, in the extreme, to preference reversal (i.e., QALY scores imply that a person prefers outcome A to outcome B, when in fact the reverse is true).[6,7] Despite these limitations, Laupacis and associates argued that "QALYs still seem to be a reasonable outcome measure for use in economic evaluations."

The question remains, How common is the QALY yardstick? Weinstein and Stason[2] describe the method for calculating QALYs as follows.

> The first approaches to this problem fall under the rubric of "health status indexes." A health status index is essentially a weighting scheme: each definable health status, ranging from death to coma to varying degrees of disability and discomfort to full health, and accounting for age differences, is assigned a weight from zero to one, and the number of years spent at a given health status, Y_s, is

multiplied by the corresponding weight, λ_s, to yield a number, $\lambda_s Y_s$, that might be thought of as an equivalent number of years with full health — a number of quality-adjusted life-years (QALYs). The source of these weights is ultimately subjective.

Laupacis and associates' approach to outcome measurement was consistent with this description. Moreover, they accepted the notion that the source of the quality weights (values of λ_s) can vary, which suggests that weights can be derived with any general (as opposed to disease-specific) measure of health-related quality of life, the values being "derived from asking either experts or patients."[1] Who the "experts" are in the subjective assessment of preferences among health states was not explained, but the implication is that they are not patients.

A more fundamental concern, however, is that if the choice of measurement method is left to individual researchers, the resultant outcome measures and the associated cost–effectiveness ratios are not common yardsticks. Many methods are available to measure the quality weights,[8] and there is no reason to believe, or evidence to suggest, that they all result in similar preference scores for the same program. For example, large systematic differences in the scores obtained by the time trade-off and the standard gamble techniques have been found.[9] This is not surprising given the fundamental differences in the two approaches,[10] but both approaches appeared to be acceptable to Laupacis and associates. A more recent study found poor correlations among six methods of assessing patients' preferences (used to measure the weights for QALYs).[11] Discrepancies among indices were particularly noticeable when data were evaluated at the individual level: many patients reported a high level of well-being according to one index and a low level according to another. This difference caused substantial variability in the calculated cost–effectiveness ratios. Because the method of measuring weights for the QALY calculation is subjective, doubt is cast on the interpretation of *comparisons* of cost/QALY among different interventions (i.e., a QALY is not a QALY . . .).

Laupacis and associates were aware of the problem of variability.

The calculated cost/QALY can vary considerably depending on the techniques used. Also, reasonable sensitivity analyses may change the cost-effectiveness of an intervention greatly. Thus, we felt that narrowing the cost/QALY ranges of the various levels any further was not justified on the basis of currently available empirical evidence and analytical techniques.

In other words, they believed that the range of the grades was large enough to prevent situations in which the choice of different techniques would result in different grades. This implies that the cost/QALY of a project is never near the threshold value, but this is an empiric question. Examples of cost–effectiveness ratios that are sensitive to alternative ways of measuring the outcome of an intervention have been reported.[12,13] In instances in which the choice of measurement technique determines the grading of a program, the guidelines do not provide a unique mechanism to determine the correct grading.

An alternative approach

We suggest an approach that maximizes the impact of health care resources on the community's health-related well-being at any given level of resources. We do not address the issue of how much should be spent on health care; rather, we focus on the level of resources allocated to health care as determined by political considerations and identify a method for allocating these resources in the most efficient way among different health care interventions. Before dealing with the decision rule, we identify a measure of the community's health-related well-being.

To simplify, we follow the QALY assumption that the community's health-related well-being is the sum of individuals' health-related well-being,[14] although we acknowledge that this may not represent the actual social welfare function. We believe that the measure of outcome at the individual level should fully represent preferences in situations of uncertainty, because health care decisions are made in an uncertain environment, at both the individual and the community level.[6,15] Hence, we suggest the use of healthy-years equivalent (HYE)[16] as a nonmonetary measure of outcome, or the willingness to pay (WTP) as a preference-based monetary measure.[17]

The advantage of both measures over the QALY is that they stem directly from the theoretic foundations of economics.[16,17] The HYE is a more recent measure than the WTP that has not yet been used in a clinical trial. But because its use involves classic tools of utility measurement (e.g., the standard gamble method), which are well developed and widely used in clinical trials, we do not foresee any problems in empirical applications. The method suggested for measuring the HYE has been tested empirically, with satisfactory results.[16] The WTP is a well-established measure of outcome that has been used in many studies.[17,18]

Neither the HYE nor the WTP is problem free.[6,16–18] Unlike the QALY, however, both methods have theoretic foundations and employ fewer and weaker assumptions about individuals' preferences

than the QALY measure. Hence, the HYE and the WTP allow individuals to reveal their true preferences better.

The health-related well-being of the community is calculated by summing HYE or WTP scores for all individuals. As mentioned previously, this assumes a particular form of the social welfare function by the implicit equal weight given to the scores of different individuals or groups.[17,19] Existing methods of QALY measurement have been shown to be inconsistent with the explicit equity statements made by researchers.[19] Methods for incorporating a chosen equity criteria into the measurement-of-outcomes method, whether QALY or HYE, have been suggested.[19] When the WTP is used a specific equity criterion is assumed.[17]

Once we have an acceptable, nonambiguous measure of outcome we must consider how to determine whether a new technology should be adopted in order to maximize clinical benefits from scarce resources (i.e., an economic perspective). The first step is to measure and calculate the total costs and consequences (e.g., in HYEs) of implementing such a program for all sizes of programs under consideration. *For a given budget*, a necessary condition for implementation would be that at least one existing program (or a combination of existing ones) will, if eliminated, (a) generate resources sufficient to provide for the new program and (b) reduce the community's health-related well-being (lost HYEs) by less than the incremental gain in the community's health-related well-being (gained HYEs) as a result of the new technology. Such a reallocation of health care resources can be seen as a "step in the right direction," because it results in an overall increase in the community's health-related well-being, without any additional resources.

It is only a step in the right direction, however, because there might be more than one program (or set of programs) that satisfy these conditions. To ensure that reallocation of resources maximizes the objective (the community's health-related well-being) we need to identify all potential programs (or combinations of programs) that meet criteria (a) and (b) and then identify the one whose elimination will reduce the community's health-related well-being the least. This results in the *maximum* possible net gain — the difference between HYEs lost and gained — in the community's health-related well-being.

The mathematical techniques to handle such an optimization problem exist, have previously been proposed for health care allocation issues[4] and are being used in other economics applications. However, the process of identifying all existing programs (or combinations of existing ones) that are eligible for substitution is not simple. We therefore recommend that, although in the long run the health care system should develop the capacity to allocate resources optimally to maximize the impact on community health-related well-being, in the short run the "step-in-the-right-direction" rule can be used to determine the technologies that should be adopted (those for which a substitution meeting the necessary conditions has to be identified) and technologies that should not (those for which no substitution can be found). This method does not result in budget increases and hence avoids making assumptions about the superiority of health care over other programs. Indeed, it coincides with the goal of maximizing the community's health-related well-being at a given level of resources.

Summary

The guidelines proposed by Laupacis and associates do not stem from economic theory and are a prescription for uncontrolled growth in health care expenditure. In particular, cost-effectiveness ratios provide information relevant to allocation decisions only in very special circumstances that do not usually apply in practice.[3] When two interventions are compared a positive cost-effectiveness ratio (the common case) can tell us, at best, what additional costs will be incurred to generate the additional outcomes. From an economic perspective the information required to determine the attractiveness of a new technology is different: the source of the additional resource requirements must be identified and the opportunity cost of their redeployment estimated.

Because the cost-effectiveness ratio (cost/QALY) is sensitive to the method chosen to calculate QALYs, guidelines that do not specify (or justify) the appropriate method for calculating outcomes are unlikely to produce comparable results (or common yardsticks). In a health care system such as Canada's in which there is always pressure to introduce more effective technology, even if it is more costly, there is a risk of using such noncomparable data to justify adoption of particular technologies.

The method of technology evaluation proposed by us is consistent with the stated goal of maximizing the community's health-related well-being for a given level of resources allocated to health care and ensures that new technologies are adopted only if this adoption represents an improvement in resource allocation.

References

1. Laupacis A, Feeny D, Detsky AS et al: How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Can Med Assoc J* 1992; 146: 473-481
2. Weinstein MC, Stason WB: Foundation of cost-effectiveness

analysis for health and medical practices. *N Engl J Med* 1977; 296: 716-721
3. Birch S, Gafni A: Cost effectiveness/utility analyses: Do current decision rules lead us to where we want to be? *J Health Econ* 1992; 11: 279-296
4. Birch S, Donaldson C: Applications of cost-benefit analysis to health care: departures from welfare economic theory. *J Health Econ* 1987; 6: 211-225
5. Loomes G, McKenzie L: The use of QALYs in health care decision-making. *Soc Sci Med* 1989; 28: 299-308
6. Mehrez A, Gafni A: Quality-adjusted life years, utility theory and healthy years equivalents. *Med Decis Making* 1989; 9: 142-149
7. Idem: Preference based outcome measures for economic evaluation of drug interventions: quality-adjusted life-years (QALYs) versus healthy years equivalents (HYEs). *Pharmacol Econ* 1992; 1: 338-345
8. Spilker B, Melinek FR, Johnston KA et al: Quality of life bibliography and indexes. *Med Care* 1990; 28 (suppl): 1-77
9. Read JL, Quinn RJ, Berwick DM et al: Preference for health outcomes: comparison of assessment methods. *Med Decis Making* 1984; 4: 315-329
10. Mehrez A, Gafni A: Evaluating health related quality of life: an indifference curve interpretation for the time trade-off technique. *Soc Sci Med* 1990; 31: 1281-1283
11. Hornberger JC, Redelmeier DA, Peterson J: Variability among methods to assess patients' well-being and consequent effect on a cost-effectiveness analysis. *J Clin Epidemiol* 1992; 45: 505-512
12. Gafni A, Zylak CJ: Ionic versus nonionic contrast media: A burden or a bargain? *Can Med Assoc J* 1990; 143: 475-478
13. Gafni A: Considerations of the measures useful for quantifying adverse effects of unnecessary hypertension drug therapy. *Clin Invest Med* 1991; 14: 266-270
14. Wagstaff A: QALYs and the equity-efficiency tradeoff. *J Health Econ* 1991; 10: 21-41
15. Ben Zion U, Gafni A: Evaluation of public investment in health care: Is the risk irrelevant? *J Health Econ* 1983; 2: 161-165
16. Mehrez A, Gafni A: Healthy years equivalent: how to measure them using the standard gamble approach. *Med Decis Making* 1991; 11: 140-146
17. Gafni A: Using willingness-to-pay as a measure of benefits: What is the relevant question to ask in the context of public decision making about health care problems? *Med Care* 1991; 29: 1246-1252
18. Johannesson M, Johansson PO, Jonsson B: Economic evaluation of drug therapy: a review of the contingent valuation method. *Pharmaco Econ* 1992; 1: 325-337
19. Gafni A, Birch S: Equity considerations in utility based measures of health outcomes in economic appraisals: an adjustment algorithm. *J Health Econ* 1991; 10: 329-342

Conferences
continued from page 899

May 12-15, 1993: Joint Conference — 39th Conference of the Canadian Organization of Medical Physicists and 19th Conference of the Canadian Medical and Biological Engineering Society
Ottawa
Dr. Paul C. Johns, assistant professor of physics, Carleton University, Herzberg Laboratories, Ottawa, ON K1S 5B6; tel (613) 788-2600, ext. 4317, fax (613) 788-4061

May 12-15, 1993: 2nd International Conference on Stroke (held under the auspices of the World Federation of Neurology)
Geneva, Switzerland
Official language: English
2nd International Conference on Stroke, c/o Kuoni Travel Ltd., Incoming Services, PO Box 1731, 7, rue de Berne, CH-1211, Geneva 1, Switzerland; tel 011-41-22-732-0888, fax 011-41-22-731-5078

May 13-14, 1993: Ontario Gerontology Association 12th Annual Conference — the Year 2000: Are We Ready?
Toronto
Debby Vigoda, Ontario Gerontology Association, 7777 Keele St., 2nd flr., Concord, ON L4K 1Y7; tel (416) 660-1056, fax (416) 660-7450

May 14, 1993: Medical Clinic Day — a Geriatric Saga
North York, Ont.
Sybil Gilinsky, Continuing Education Department, Baycrest Centre for Geriatric Care, 3560 Bathurst St., North York, ON M6A 2E1; tel (416) 789-5131, ext. 2365

May 14-15, 1993: Atlantic Provinces Ophthalmological Society Annual Meeting
Halifax
Dr. Paul A. Price, Department of Ophthalmology, St. Martha's Regional Hospital, Antigonish, NS B2G 2G5; tel (902) 863-6411, fax (902) 863-6006

May 14-15, 1993: Ischemic Heart Disease, Exercise and Related Topics — 5th International Symposium
Toronto
Study credits available.
Symposium coordinator, Toronto Rehabilitation Centre, 347 Rumsey Rd., Toronto, ON M4G 1R7; tel (416) 425-1117, fax (416) 425-0301

May 15, 1993: Education in Occupational Medicine
Hamilton, Ont.
Dr. David Muir, McMaster University Medical Centre, Rm. 3H50, 1200 Main St. W, Hamilton, ON L8N 3Z5; tel (416) 525-9140, ext. 2332, fax (416) 528-8860

May 15-16, 1993: 6th Annual GP Psychotherapy Conference — Bringing Psychotherapy to Life by Bridging the Gap Between Medicine and Psychotherapy
Toronto
Dr. Dianne McGibbon, 3 Gardenvale Rd., Toronto, ON M8Z 4B8; tel (416) 239-4644

May 15-18, 1993: Power of Partnership — Annual Convention of the Catholic Health Association of Canada
Ottawa
Freda Fraser, director of communications, Catholic Health Association of Canada, 1247 Kilborn Pl., Ottawa, ON K1H 6K9; tel (613) 731-7148, fax (613) 731-7797

continued on page 938