

REFERENCES

1. *In re Quinlan*, 70 N.J. 10, 355 A. 2d 647, cert. denied sub nom., *Garger v. New Jersey*, 429 U.S. 922 (1976).
2. *In re Storar*, 52 N.Y. 2d 363, 420 N.E. 2d 64, cert. denied, 454 U.S. 858 (1981).
3. *In re Westchester County Medical Center on behalf of O'Connor*, 581 N.E. 2d 607 (N.Y. 1988).
4. *Cruzan v. Harmon*, 760 S.W. 2d 408 (Mo. 1988) (en banc).
5. *Cruzan v. Director, Missouri Dept. of Health*, 110 S.Ct. 2841 (1990).
6. *In re Jobes*, 108 N.J. 894, 529 A. 2d 434 (1987).
7. Annas GJ, Arnold B, Aroskar M, et al. Bioethicists' statement on the U.S. Supreme Court's *Cruzan* decision. *N Engl J Med* 1990; 323:686-7.
8. *Ohio v. Akron Center for Reproductive Health*, 110 S.Ct. 2972 (1990).
9. Legal Advisors Committee, Concern for Dying. The right to refuse treatment: a model act. *Am J Public Health* 1983; 73:918-21.

THE NEW TECHNOLOGY ASSESSMENT

TECHNOLOGY assessment is not a new phenomenon in medicine. The most able and conscientious physicians have always sought to understand the effects of the interventions they apply.¹ With the development of clinical research, attempts to establish safety and efficacy became more systematic and scientific, culminating in the crown jewel of traditional technology assessment, the randomized clinical trial.

The determination of safety and efficacy remains an essential element of technology assessments, but in recent years such assessments have broadened in scope. They now encompass the measurement of effectiveness, considerations of quality of life and patients' preferences,^{2,3} and especially the evaluation of costs and benefits.⁴ Although such assessments were published as early as the mid-1970s,^{5,6} widespread interest in the new type of technology assessment has flourished only recently, stimulated by concern about the quality, effectiveness, and escalating cost of health care.^{7,8} When Congress passed and the President signed the Omnibus Budget Reconciliation Act of 1989, the prospect of more influential, better-funded technology assessments was created; the act called for almost \$600 million to be spent over the next five years to support research on the outcomes, effectiveness, and appropriateness of medical care. Private foundations also believe that technology assessment can make important contributions to American health care, and they have allocated tens of millions of dollars to this activity.

Admiration for the new technology assessment is not universal. Many practicing physicians believe it will further erode their ability to practice as they deem best; similarly, medical researchers, pharmaceutical manufacturers, and producers of medical devices fear that it will inhibit the development and diffusion of new forms of technology. These fears may be justified, but the new technology assessment is here to stay and will probably have a large effect on medical practice. It is therefore imperative for physicians and the producers of health care technology — along with others who aspire to find the best uses for health dollars — to try to understand and improve it. To do so, it

is necessary to know how the new technology assessment differs from the old type, what is special about it in medicine, what features make an assessment credible, and when one should be performed. Most important, we must recognize what an assessment can and cannot do.

WHAT IS NEW ABOUT THE NEW TECHNOLOGY ASSESSMENT?

The most striking differences between the new and the old forms of technology assessment arise from a broadening of perspective. The old form emphasized the biomedical perspective — that is, the safety and efficacy of an intervention. The new technology assessment, with its broader perspective, is usually conducted by different researchers who apply different methods and seek different data. Because it evaluates a wider spectrum of consequences of health interventions, the new technology assessment is more challenging, more complex, more controversial, and potentially more useful than the old one.

The distinguishing features of the new technology assessment can be discerned in its approach to specific forms of technology. Consider, for instance, a new test for cystic fibrosis⁹ that detects a mutation in the recently identified gene associated with the disease.¹⁰ This test is particularly suited to analysis with the new technology assessment because technical performance is only one of its important characteristics. It is necessary, but not sufficient, to confirm that the test detects the mutation. There are additional questions. How sensitive and specific is the test in various populations? What will families do after they learn that a fetus has tested positive for cystic fibrosis? How do the costs of screening for cystic fibrosis compare with the benefits? What are the ethical implications of using the new test to screen the general population, when even small false positive rates will lead to the abortion of many normal fetuses? Like its predecessor, the new technology assessment explores biomedical characteristics, but it also grapples with broader issues such as these.

As the perspective of technology assessment has changed, so has the identity of the assessors. Initially, assessments were carried out by the scientists and physicians who developed a form of technology or used it in caring for their patients. As formal trials of interventions became more common, biostatisticians and other specialists in the design and performance of clinical trials came to have important roles. With its increased emphasis on costs, quality of life, and patient satisfaction, the new technology assessment continues to draw on the traditional areas of expertise, but the skills of collaborating researchers trained in such disciplines as economics, epidemiology, operations research, and psychometry have been added. These researchers use a wide variety of sources and methods to carry out their evaluations.

There is also a difference in the audience for technology assessment — that is, its consumers. Formerly,

the results were intended primarily for physicians and were communicated in professional journals and at scientific meetings. One goal of the new technology assessment is to influence physicians, but assessments are also frequently addressed to (and undertaken at the behest of) health insurance companies, employers, and government agencies and policy makers. The public, which often learns of the results of assessments through the media, is also an interested and sometimes the intended audience.

WHAT IS SPECIAL ABOUT TECHNOLOGY ASSESSMENT IN MEDICINE?

Technology assessment in medicine differs from that in most industries in at least three ways. First, the assessment process is technically complex; second, health insurance coverage reduces the patient's incentive to evaluate the cost effectiveness of forms of technology; and third, the health care industry is highly fragmented, with each company or health care provider controlling only a small share of the total national market.

The complexity of most forms of technology in health care can be appreciated by contrasting them with innovations such as more fuel-efficient automobile engines. The average driver may understand virtually nothing about internal combustion but can readily determine how many more miles per gallon a new engine delivers. In contrast, even an expertly designed clinical trial carried out by experienced investigators may not yield clear-cut, quantitative answers about the effectiveness of a new diagnostic or therapeutic intervention.

Furthermore, if the new engine does save gasoline but increases the cost of the car, the buyer has both the incentive and the ability to determine whether the saving in gasoline is worth the extra cost. Knowing that buyers will make this calculation, automobile manufacturers make it first. Thus, innovations that do not promise to be cost effective are weeded out at an early stage of development. In contrast, under conventional programs of health insurance and reimbursement, "buyers" bear only a fraction of the costs, so health care innovators do not have the incentive to carry out such assessments.

The contrast with automobile manufacturing also obtains with respect to the scale of activity. General Motors, Ford, Toyota, and the like each control a substantial share of the market. When they invest in technology assessment, they stand to benefit directly from the results. Individual physicians, however, lack the incentive or ability to assess new forms of medical technology; any such assessment is at least in part a "public good." Even the largest insurance companies each account for only a small percentage of the health care market. They are therefore understandably reluctant to pay for large-scale assessments that will benefit all. As a result, the new technology assessment in medicine is funded principally by government

agencies, private foundations, and to a far lesser extent, private third-party payers and managed care plans.

Although support for technology assessment from government and foundations is not unique to medicine, the extent of reliance on these funding sources is unusual. But there are alternatives. Health insurance companies might well consider the model provided by the electric-power industry, in which a small portion of each public-utility bill goes to fund the Electric Power Research Institute, a major research organization devoted to the development and especially the assessment of forms of technology related to that industry. If all private health insurers devoted just one penny of each \$10 of premium income to technology assessment, \$200 million per year would become available for this purpose. Research on this scale, if it leads to changes in practice, could result in annual savings of billions of dollars in health care expenditures.

WHAT MAKES A GOOD ASSESSMENT?

As the goals of technology assessment in health care have grown more ambitious, its methods have become more sophisticated.¹¹⁻¹⁴ They make it possible to answer broad questions even when the data are incomplete, and they can provide greater assurance than simpler methods that the conclusions of a study follow from its data and assumptions. As the studies become more complex, however, the proportion of the audience that feels competent to evaluate the assumptions and methods shrinks, and the potential for misleading readers grows. Thus, many consumers of the new technology assessment rely on more than their own reading of a study for assurances that the analysis is valid. The peer-review process in journal publication, formal processes of committee and organizational review (such as the approval process used by the American College of Physicians in its Clinical Efficacy Assessment Project¹⁵), and reviews by groups with opposing interests are all helpful ways for practicing physicians and other decision makers to judge the validity of the conclusions.

Despite these checks, there are two reasons to fear that a complicated analysis may not be valid. First, the investigators may have a conflict of interest that did not become known during the review process. Second, the authors and reviewers may fail to detect errors because of shortcomings in the data, the analysis, or the interpretation of the results. As in established fields of biomedical research, to safeguard against such errors, the evaluation of forms of medical technology depends in part on the integrity, objectivity, carefulness, and expertise of the researchers who conduct it.

The validity of a study, however, depends most on how it is conducted, rather than by whom. The authors of technology assessments often minimize the effect of bias and inadvertent errors by follow-

ing steps similar to those expected of any scientific study. They make available the data used for the analysis, either including it in the report or furnishing it on request. For example, assessments that employ a meta-analysis of randomized, controlled trials¹⁶⁻¹⁹ often display in detail the data and the criteria for the inclusion of the data sources. In addition, the investigators describe clearly and prominently the assumptions used in the analysis and perform a sensitivity analysis to show the consequences of altering uncertain assumptions. Studies that follow these steps allow knowledgeable readers to reject or confirm the conclusions in a way rarely feasible with laboratory research. Impartiality, critical judgment, expertise in the form of technology discussed and its alternatives, and command of the relevant methods distinguish the best assessors; the best evaluations are complete, explicit, and balanced.

STAGES AND TIMING OF TECHNOLOGY ASSESSMENT

Technology assessment usually proceeds in three stages, with a progressively wider scope. The first stage focuses on technical characteristics, such as the *in vitro* antibacterial activity of an antibiotic, the degree of resolution of a scanner, or the short-term safety of a drug. In the second stage, aspects pertaining to the efficacy of a diagnostic or therapeutic technique are usually investigated. Traditional randomized, controlled clinical trials often fall into this category. Sometimes these studies use intermediate or surrogate end points, because the studies are too small or require too much time to detect statistically significant effects on the end points that matter to patients. Studies of diagnostic techniques in this stage often emphasize sensitivity and specificity. In some respects, the comparison of magnetic resonance imaging (MRI) and ultrasonography for the staging of prostate cancer by Rifkin and colleagues, reported in this issue of the *Journal*,²⁰ falls into this category.

The third stage of technology assessment, which distinguishes the new technology assessment from the old, focuses on the final outcomes of interest — the clinical, economic, and social end points. Rifkin and colleagues implicitly recognize the importance of this stage. On finding that the positive predictive value of MRI scanning for nodal involvement is disappointingly low, they report that “Whether this percentage is too low to indicate a treatment strategy is unclear and would depend on a weighing of the side effects of biopsy against the benefits of the information thus obtained.” By making this statement, the authors imply that the next logical step is part of a third-stage assessment: a comprehensive analysis of how the two techniques should be used in the care of patients with prostate cancer. Similarly, third-stage assessments of the use of thrombolytic therapy for myocardial infarction would measure the changes in mortality and disability that result from treatment and relate them to the cost of therapy. Third-stage assessments of chole-

sterol-lowering medications would evaluate the reduction in heart attacks, angina, and mortality that results from treatment. In addition, some third-stage assessments are beginning to incorporate information about patients’ preferences regarding the multiple effects of health interventions.²¹

Limitations on the availability of data usually mandate that technology assessment proceeds sequentially from the first stage to the third, but that is not always the case, nor should it be. The developer of a new intervention needs to decide at an early stage whether there is any reasonable expectation that it will substantially alter health outcomes. A new test to detect an untreatable condition, for example, might not affect outcomes even if it was highly sensitive and specific. The developer could perform a third-stage assessment with hypothetical data before exact figures on test performance were available; even a preliminary analysis might reveal that the new diagnostic technique would be unlikely to improve health. Phelps and Mushlin have shown that early assessments can help estimate the value of a technology and determine whether it merits further investigation.¹⁴ Any stage of a technology assessment could be performed before all the data pertinent to its evaluation become available, although lack of data may limit its value.

Like the traditional technology assessment, the new type must reassess techniques periodically as new data about safety and efficacy become available, as the technique is used for new indications, or as the intervention itself changes.²² For example, the declining level of exposure to radiation needed to perform mammography makes that procedure more attractive as a screening test. Reevaluation is more important for the new technology assessment, however, since unlike the traditional assessment, the new form considers changes in other factors, such as cost, that make a technique more or less desirable. The Health Care Financing Administration rejected the practice of making extra payments to hospitals that used recombinant tissue-type plasminogen activator (TPA) to treat patients with a myocardial infarction. Advocates of TPA claimed that it was more effective, or safer, than the chief alternative (streptokinase), but at the time of the decision TPA cost nearly 10 times as much as the least expensive alternative. Many observers thought that if TPA had been priced lower, it would have been selected as the thrombolytic agent of choice. A change in the price of a thrombolytic agent or any other form of technology changes its cost effectiveness — a consideration that is fundamental to the new but not the traditional technology assessment.

WHAT TECHNOLOGY ASSESSMENT CAN AND CANNOT DO

Efforts to contain costs and monitor the quality of medical care rely increasingly on formal technology assessment.^{23,24} By helping to define appropriate care, such assessment is often instrumental in the develop-

ment of health policy and the setting of reimbursement guidelines. An unequivocal assessment, whether in a randomized, controlled clinical trial or a comprehensive "new" assessment, can promote the dissemination or rejection of a form of technology. The possibility that much modern, costly medical care is "flat of the curve medicine"²⁵ (providing few additional health benefits) is a key reason for engaging in technology assessment. When assessment catalyzes the abandonment of interventions whose risks exceed their health benefits, it can lower the costs of health care without sacrificing its quality. For several reasons, though, clinical and policy decisions about health practices must usually depend on more than the results of a technology assessment. The poor quality of the study is sometimes the reason; an assessment may be so flawed by gaps in the data or shortcomings of the analytical methods that it cannot answer the questions posed. Traditional technology assessment also confronted imperfect data, but because its goals were narrower, it usually required less information and used more straightforward methods for analyzing the data.

To readers unfamiliar with the general approach, even rigorous and careful assessments can appear untrustworthy. Such readers may be skeptical because the new technology assessment uses novel methods and data from diverse sources. Many physicians believe that only a randomized, controlled clinical trial can provide definitive measures of the effectiveness of an intervention. The new technology assessment often includes data from randomized trials, but is not limited to them. Randomized trials usually cannot provide all the information needed for an assessment; physicians may need to decide whether to adopt a form of technology long before the results of such a trial become available, and in a trial it is not always feasible to measure all the relevant end points. Furthermore, trials are frequently performed in clinical settings different from those in which the interventions are most likely to be applied. For many interventions, ethical or logistical difficulties may preclude conducting a randomized trial. Instead of ignoring questions that cannot be answered in the rigorous framework of a trial, the new technology assessment seeks answers by applying such methods as multivariate analysis of nonexperimental data,^{26,27} meta-analysis, decision analysis, and economic modeling. Although both the data and the techniques are appropriate, they often make practicing physicians and biomedical researchers uncomfortable, thus limiting the acceptance of the findings.

Demonstrating scientific credibility is not the only challenge the new technology assessment faces. Even otherwise well-executed assessments may not resolve whether a form of technology is worth using, because they stop short of assessing what a change in health is worth in dollar terms. In cases in which technology assessment reveals that an intervention provides no net benefit, the assessment may be decisive. But em-

ployers and third-party payers have not promoted cost-containment programs merely to avoid paying for useless interventions. They also seek to limit the application of costly interventions that have small, albeit positive, effects on health. From a payer's point of view, an intervention must be cost effective, meaning that its health benefits are commensurate with the benefits from interventions of equal or lower cost. Although technology assessment often includes an analysis of cost effectiveness, it cannot tell us how much we should be willing to pay for a given health effect. That would require explicit value judgments, which are eschewed in both the old type of technology assessment and much of the new.

Sometimes the results of a technology assessment are used, but not in the way the authors intended. Sponsors and third-party payers often expect to use the results of an assessment to develop practice guidelines or reimbursement policies. For them, the best outcome of a study is a simple yes-or-no answer to the question, Is this intervention worthwhile? But most physicians know that few promising forms of technology are simply good or bad; their value usually depends on the clinical setting and the preferences of the patient. The benefits of immediate surgery for benign prostatic hypertrophy, for example, depend on the patient's attitude toward different states of health.²¹ It will not be easy to develop reimbursement policies that take such factors into account. Collecting detailed information about each patient is expensive, and physicians and patients can usually circumvent policies based on subjective criteria. Payers may assume with some justification that even if they agree to pay for a treatment only if it is applied to the subgroup of patients for whom it is highly cost effective, the treatment will often be given to patients for whom it is much less cost effective. Thus, studies that demonstrate a substantial benefit for a small group of patients but minimal benefit or none for a larger group may convince payers who cannot monitor or control the use of a form of technology to deny coverage altogether or to impose seemingly arbitrary rules for its use, such as age limits for hemodialysis.

Payers use reimbursement policies to contain health care expenditures because they know that clinical information is only one of many factors influencing medical practice. Characteristics of patients are only partly responsible for the widespread variations in rates of operations and other medical procedures in apparently similar populations.²⁸⁻³⁰ It seems unlikely that differences in knowledge account for these variations, which reflect the behavior of both physicians and patients. Undoubtedly, medical education and training, the organization of practice, the legal system, and peer review influence the behavior of physicians. Education, the family, the workplace, and community values can have important effects on the behavior of patients. But few influences that are amenable to policy intervention are as important as the economic incentives that both physicians and patients

face. These incentives often promote the overuse and inefficiency in health practices that are responsible for much of the variation in use. Although technology assessment is helpful for setting reimbursement policies, voluntary guidelines based on it are insufficient to impel changes in the behavior of physicians.³¹ To achieve such changes, private and public policy makers must understand how incentives shape behavior and must act on that understanding.

IMPLICATIONS FOR PRACTICING PHYSICIANS

The effects of the new technology assessment, like those of any tool, depend on the way it is used. To physicians, the new method of assessment represents both a threat and an opportunity. Government and private payers may use it to micromanage health care and interfere further with the physician-patient relationship. But it can provide physicians with critical information with which to evaluate and improve their clinical practices. To influence the future of health care, physicians need to recognize the legitimacy of the broad goals addressed by the new technology assessment. And although physicians should examine assessments with a critical eye, they should not dismiss them out of hand, even when the assessments are used to constrain practice. The new technology assessment explicitly incorporates considerations that clinicians have long claimed are integral to making decisions about the care of individual patients — the circumstances, needs, and values unique to each. In an era when government, insurance companies, and employers insist on cost containment, the new technology assessment can be the most powerful tool physicians have to protect the interests of their patients.

Stanford University
Stanford, CA 94305

VICTOR R. FUCHS, PH.D.
ALAN M. GARBER, M.D., PH.D.

We are indebted to Douglas Owens, M.D., Byron W. Brown, Ph.D., and Mark Hlatky, M.D., for helpful comments.

Dr. Garber is a Henry J. Kaiser Family Foundation Faculty Scholar in General Internal Medicine.

Address reprint requests to Dr. Garber at the National Bureau of Economic Research, 204 Junipero Serra Blvd., Stanford, CA 94305.

REFERENCES

- Major RH. A history of medicine. Springfield, Ill.: Charles C Thomas, 1954.
- Fowler FJ Jr, Wennberg JE, Timothy RP, Barry MJ, Mulley AG Jr, Hanley D. Symptom status and quality of life following prostatectomy. *JAMA* 1988; 259:3018-22.
- McNeil BJ, Weichselbaum R, Pauker SG. Fallacy of the five-year survival in lung cancer. *N Engl J Med* 1978; 299:1397-401.
- Office of Technology Assessment, U.S. Congress. The implications of cost-effectiveness analysis of medical technology: background paper. Washington, D.C.: Government Printing Office, 1980.
- Weinstein MC, Stason WB. Hypertension: a policy perspective. Cambridge, Mass.: Harvard University Press, 1976.
- Bunker JP, Barnes BA, Mosteller F. Costs, risks, and benefits of surgery. New York: Oxford University Press, 1977.
- Roper WL, Winkenwerder W, Hackbarth GM, Krakauer H. Effectiveness in health care: an initiative to evaluate and improve medical practice. *N Engl J Med* 1988; 319:1197-202.
- Committee for Evaluating Medical Technologies in Clinical Use, Division of Health Sciences Policy, Division of Health Promotion and Disease Prevention, Institute of Medicine. Assessing medical technologies. Washington, D.C.: National Academy Press, 1985.
- Lemna WK, Feldman GL, Kerem B-s, et al. Mutation analysis for heterozygote detection and the prenatal diagnosis of cystic fibrosis. *N Engl J Med* 1990; 322:291-6.
- Kerem B-s, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. *Science* 1989; 245:1073-80.
- Eddy DM. Screening for cancer: theory, analysis, and design. Englewood Cliffs, N.J.: Prentice-Hall, 1980.
- Idem*. The confidence profile method: a Bayesian method for assessing health technologies. *Oper Res* 1989; 37:210-28.
- Shachter RD, Eddy DM, Hasselblad V. An influence diagram approach to medical technology assessment. In: Oliver RM, Smith JQ, eds. Influence diagrams, belief nets and decision analysis. London: Wiley, 1990.
- Phelps CE, Mushlin AI. Focusing technology assessment using medical decision theory. *Med Decis Making* 1988; 8:279-89.
- Schwartz JS, Ball JR, Moser RH. Safety, efficacy, and effectiveness of clinical practices: a new initiative. *Ann Intern Med* 1982; 96:246-7.
- Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987; 316:450-5.
- L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987; 107:224-33.
- Littenberg B, Garber AM, Sox HC Jr. Screening for hypertension. *Ann Intern Med* 1990; 112:192-202.
- Weinstein MC. Cost-effective priorities for cancer prevention. *Science* 1983; 221:17-23.
- Rifkin MD, Zerhouni EA, Gatsonis CA, et al. Comparison of magnetic resonance imaging and ultrasonography in staging early prostate cancer: results of a multi-institutional cooperative trial. *N Engl J Med* 1990; 323:621-6.
- Barry MJ, Mulley AG Jr, Fowler FJ, Wennberg JW. Watchful waiting vs. immediate transurethral resection for symptomatic prostatism: the importance of patients' preferences. *JAMA* 1988; 259:3010-7.
- Banta HD, Thacker SB. The case for reassessment of health care technology: once is not enough. *JAMA* 1990; 264:235-40.
- Relman AS. Assessment and accountability: the third revolution in medical care. *N Engl J Med* 1988; 319:1220-2.
- Russell LB. Some of the tough decisions required by a national health plan. *Science* 1989; 246:892-6.
- Fuchs VR. Health care and the United States economic system: an essay in abnormal physiology. *Milbank Mem Fund Q* 1972; 50:211-37.
- Sechrest L, Perrin E, Bunker J, eds. Research methodology: strengthening causal interpretations of nonexperimental data. Washington, D.C.: Government Printing Office, 1990. (DHHS publication no. (PHS) 90-3454.)
- Moses LE, Brown BW Jr. Experiences with evaluating the safety and efficacy of medical technologies. *Annu Rev Public Health* 1984; 5:267-92.
- Wennberg J, Gittelsohn A. Small area variations in health care delivery: a population-based health information system can guide planning and regulatory decision-making. *Science* 1973; 182:1102-8.
- Leape LL, Park RE, Solomon DH, Chassin MR, Koseoff J, Brook RH. Does inappropriate use explain small-area variations in the use of health care services? *JAMA* 1990; 263:669-72.
- Chassin MR, Koseoff J, Park RE, et al. Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. *JAMA* 1987; 258:2533-7.
- Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. *N Engl J Med* 1989; 321:1306-11.

CORRESPONDENCE



NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY OF PLASMA FOR THE DETECTION OF MALIGNANT TUMORS

To the Editor: When matched with biopsy and operative pathological reports, our results,¹⁻³ as well as those from certain laboratories,⁴⁻⁶ demonstrate that our nuclear magnetic resonance (NMR)