

Cost-Effectiveness And Evidence Evaluation As Criteria For Coverage Policy

Cost-effectiveness analysis could shift from being an academic curiosity to an essential tool for health care decision making.

by **Alan M. Garber**

ABSTRACT: Private health plans and government health insurance programs in the United States base their coverage decisions on evidence criteria, rather than explicit cost-effectiveness criteria. As health spending continues to grow rapidly, however, approaches to coverage policy that ignore costs fail to meet the needs of consumers, employers, health plans, and federal and state governments. I describe the role of evidence-based criteria in formal coverage decision making and contrast the ways that these criteria differ from cost-effectiveness criteria. Finally, I discuss options for incorporating considerations of cost-effectiveness into coverage policy and other aspects of benefit design.

RESURGENT HEALTH SPENDING GROWTH and the continuing erosion of private health insurance have renewed U.S. debates over health care reform. Absent from these debates, however, is any systematic discussion of processes to choose the medical goods and services that health insurance should cover. Policymakers may instinctively sidestep the topic as a narrowly technical issue, to be decided by physicians and others with the patience and interest to evaluate a mass of information about medical treatments and diagnostic tests. They may also see little incentive to pursue it, knowing the political risk that comes with any public stand on coverage policy.

Their reticence is unfortunate, though, because coverage policy is so tightly linked to the affordability of health insurance, and hence the rate of uninsurance. When the cost of purchasing a private health insurance plan rises, the number of Americans with commercial health insurance falls: Employers stop offering their employees health insurance, and employees stop paying their share of premiums when their employers continue to offer insurance. Coverage policy also influences the types of medical care Americans receive, because health insurance coverage is the gateway to the availability of medical innovations. It is difficult to imagine

Alan Garber (garber@stanford.edu) is staff physician, Veterans Affairs Palo Alto Health Care System. He is also the Henry J. Kaiser Jr. Professor and professor of medicine, director of the Center for Primary Care and Outcomes Research, and director of the Center for Health Policy at Stanford University in Stanford, California.

how therapies that cost thousands of dollars per patient, such as left ventricular assist devices for severe congestive heart failure, could be adopted if health insurance did not cover them.

Although they did not arise from an explicit legislative process, de facto principles for coverage decision making have emerged. They are the product of historical practices, legal decisions, and insurance contract language. Coverage policy under both Medicare and most commercial health insurance plans is based upon a determination that a medical product has proved to be effective. That is, most explicit processes for making coverage decisions in the United States are based on evidence, not on cost-effectiveness or any other direct measure of value.¹

This paper discusses the similarities and differences in using evidence and value criteria as bases for coverage decision making. It also addresses the complementary roles of coverage policy and other approaches to limiting the cost of health insurance, such as increased cost sharing, and the implications for the design of health insurance.

Coverage Policy And The Costs Of Health Insurance

Economists have long identified “moral hazard,” the overuse of health services that occurs because the insured person bears only a fraction of the cost of covered services, as the chief cause of excessive health spending.² Moral hazard raises the level of health spending. By stimulating the development of new technologies, it also increases the rate of spending growth. The expectation that health insurance will boost demand and revenues in the future is a powerful incentive to invest in the development of new medical technologies. The high rate of innovation in medical products and services leads to better health but also higher spending.³

The simplest instrument insurers have to keep costs down is to negotiate favorable reimbursement rates with providers. Such strategies, of course, have little ability to offset the growth in spending that results from heavier use of costly services. Thus, it may be more important to control moral hazard.⁴ Their main tools for discouraging excessive use are supply-side incentives, direct utilization controls, and copayments and deductibles that expose patients to the financial consequences of their use of health care. In addition, all health insurance plans place boundaries on the products and services that qualify for reimbursement. These boundaries constitute the insurer’s coverage policy.

Coverage is described in health insurance contracts, which list entire categories of products and services that are excluded from coverage, such as cosmetic surgery, as well as categories of inclusion, such as hospitalization for medical emergencies. The contracts cannot provide detailed descriptions of every service that will and will not be covered within a category of services eligible for inclusion, so they usually state that the insurer will reimburse all “medically necessary” goods and services. The interpretation of the term “medically necessary” has varied over time and across health plans, but today it generally rests upon the application of

an evidence standard.⁵

Such standards have two critical components: a determination about whether enough evidence is available to support conclusions about the effectiveness of the intervention in question (adequacy of evidence), and a determination about what that evidence implies about effectiveness (magnitude of benefit). This approach promises to reduce waste and improve safety by avoiding payment for products and services that are likely to be harmful or of no benefit.⁶ The emphasis on high-quality evidence represents a marked change from an earlier era of medicine, when the doctor's beliefs about the value of an intervention, especially if they were widely shared, were sufficient to establish medical necessity.

Although different groups do not always reach the same conclusion about a particular technology, and the specific processes that they use to evaluate evidence vary, there has been a remarkable convergence in the acceptance of the principle that coverage determinations—and indeed, medical practice itself—should be guided by the results of rigorously designed studies, rather than expert opinion or the most common forms of practice. Particularly for processes that are intended to inform coverage decisions for large numbers of people, such as the Blue Cross Blue Shield Association's Medical Advisory Panel and the federal Medicare Coverage Advisory Committee (MCAC), these evidence processes typically place great weight on information from well-designed clinical trials.⁷ The approach to evidence evaluation is similar to approval processes used by the U.S. Food and Drug Administration (FDA) and to the evidence ratings pioneered by Canadian and U.S. task forces on preventive services.⁸

Randomized clinical trials have great influence because they are less susceptible to bias than are studies with less rigorous design. Observational studies sometimes accurately predict the results of randomized trials, particularly in areas such as the treatment of heart disease. However, observational studies in many other areas, such as cancer treatment, are highly susceptible to bias.⁹ Treatments that appeared to provide large benefits in well-designed observational studies, such as bone marrow transplantation for advanced breast cancer, were found to be ineffective in randomized trials, perhaps because the women who received more aggressive treatment in observational settings were healthier at the outset.

For health plans, rigorous evidence-based processes have a powerful appeal: It is difficult to argue that an ineffective test or treatment should be covered by insurance or even administered by a physician. Cutting waste without eliminating effective care would seem a painless way to begin to limit medical spending. But the adoption of an evidence standard does not represent solely a commitment to avoid ineffective care. Since no intervention is assumed to be effective until it has been proved effective, the burden of proof for a new medical intervention is placed on its advocates. Examining the evidence requirement from their point of view is an important step toward understanding its consequences.

The Burden Of Establishing Effectiveness

Meeting an evidence standard is costly. The drug approval process offers a view into how an evidence-based process works and what it costs. Of course, evaluations of diagnostic tests and surgical procedures differ in important respects from the evaluation of drugs for FDA approval, and drugs are a relatively small fraction of the mix of products and services that health plans cover. Nevertheless, the most critical issues faced by any group or individual seeking to demonstrate effectiveness are common to all medical interventions.

Tests of safety and effectiveness in humans are believed to be responsible for more of the cost of drug development than basic drug discovery research. In recent years, the cost of large-scale clinical trials for a successful drug, according to a study based on industry-supplied data, averaged about \$86.3 million (in 2000 dollars).¹⁰ Per patient costs of trials in the United States are estimated to fall in the range of \$10,000–\$50,000.

These high costs make it essential to keep trials as small as possible. However, a trial that enrolls too few patients will be unable to demonstrate conclusive (statistically significant) evidence of benefit. The magnitude of the intervention's health improvement, the characteristics of the patients enrolled in the trial, and the variability in their health outcomes are among the factors that determine how large a trial is needed to achieve statistical significance. Investigators can improve a study's prospects of success by enrolling participants who are likely to show the greatest benefit most quickly, or by increasing either the duration of follow-up or the number of patients studied. The importance of sample size can hardly be overstated. A pooled analysis of clinical trials conducted in the early 1990s noted that the vast majority of trials reporting negative results did not have adequate statistical power, despite large effect sizes (relative changes in the health outcome) of 25 percent or even 50 percent.¹¹

Although they cannot eliminate the risk and uncertainty inherent in evaluating a new intervention—if there were no uncertainty, there would be no reason to conduct a trial—researchers can prevent many of the pitfalls by increasing the size or duration of the trial.¹² The combination of high per patient costs and the need to have adequate sample sizes is responsible for the high costs of clinical trials.

Although these costs can be a daunting obstacle, the prospect of revenues from monopoly in the sale of the intervention is a strong enough incentive for both pharmaceutical companies and device manufacturers to fund trials.¹³ Lacking well-defined, enforceable intellectual property rights, developers of innovative care processes and medical procedures have little prospect of gaining a monopoly. They cannot expect future payments large enough to offset the cost of studies that would establish effectiveness. Perhaps that is one reason why many surgical innovations are tied to the use or implantation of a patented device (for example, left ventricular assist device, implantable cardioverter-defibrillator, or coronary stent). Interventions that do not lead to monopoly products are sometimes stud-

ied with the support of the National Institutes of Health (NIH), the Department of Veterans Affairs (VA), and other government agencies, but federal funds only support trials of a fraction of promising forms of care.

Thus, evidence-based processes, which usually build upon explicit, statistically based criteria, are subject to the important qualification that someone had to have conducted a convincing study. Because monopoly rewards are often the chief incentive to fund research, evidence standards tend to favor monopoly products over other approaches to improving health outcomes, such as a new use for a generic drug, a better diagnostic strategy, or an improvement in delivering care. A bias toward such products, in turn, has important implications for spending.¹⁴

Applying Cost-Effectiveness Analysis To Coverage Decisions

Advocates for quality improvement remind us that evidence-based processes reduce spending by discouraging the use of ineffective medical care. Cutting waste is an attractive way to cut the level of health spending, but it may not slow its rate of growth. Most innovation represents improvements in care, and it is the growth in the volume and intensity of care, not disproportionate growth in wasteful care, that drives medical spending.¹⁵ Cost-effectiveness analysis can complement strategies to eliminate waste, since it can be used to guide utilization away from procedures that produce little benefit at high cost—in other words, to improve the efficiency of health care.¹⁶

Ideally, health insurance would promote the use of cost-effective medical services. It might do so by covering only services whose cost-effectiveness ratio is equal to or less than a cutoff (threshold) value.¹⁷ Under specific assumptions, the cutoff can be inferred from individual preferences, but the limited literature on this topic has not led to a consensus about how such thresholds should be determined and used.¹⁸ For example, if the cost-effectiveness threshold were based upon a person's willingness to pay for an improvement in health, the threshold would vary from one person to another. But many proponents of using cost-effectiveness analysis for health care decision making would apply a single threshold to an entire population.

Another approach would avoid selecting a threshold cost-effectiveness ratio and would instead compare the cost-effectiveness of various widely used interventions, giving an idea of the value of the intervention relative to other familiar health interventions. The ranking of the cost-effectiveness of various interventions is presented in a "league table." This approach has been criticized in part because the tables often report results that have been obtained from studies using different, and often incompatible, methods.¹⁹ The problem is particularly severe when studies use different measures of health effects: One may use changes in life expectancy; another, changes in quality-adjusted life years (QALYs); and yet another, changes in cholesterol or blood pressure. Most importantly, although a league table ranks interventions by their cost-effectiveness ratios, it does not tell

the reader where to draw the line between acceptable and unacceptable interventions—indeed, that would be equivalent to selecting a cost-effectiveness cutoff.²⁰

Setting the cutoff at a level that would lead to the rejection of potentially life-saving procedures is controversial among those who expect that all effective care will be available to everyone. Furthermore, rigid application of a specific cutoff cost-effectiveness ratio is rarely possible—if only because effectiveness varies from one person to another—nor would it guarantee socially acceptable outcomes. Awareness of the incompleteness of the threshold as a decision criterion has led expert panels to conclude that it should be combined with other information to guide clinical and policy decisions. For example, they would consider whether other treatments are available for the disease in question. They might also modify standards to shift care toward underserved racial or ethnic groups. This is similar to the approach that Oregon adopted in its attempt to distribute Medicaid funds to a broader population of uninsured people. Oregon started with a ranking of procedures based principally on cost-effectiveness but developed a very different list after extensive public discussion.²¹

Once the details of such a process are determined, how do the resulting choices differ from those based on an evidence-based approach? We begin by asking which interventions that are highly effective for their cost will readily pass an evidence standard, and which will not.

Both methods are likely to “pass” an intervention that is inexpensive and highly effective. A relatively small sample size or a short-duration trial, or both, would be sufficient to establish a statistically significant benefit. If the intervention is extremely expensive, it will pass an evidence but not a cost-effectiveness criterion.²²

Beyond these general points, the higher costs of establishing effectiveness in a trial—largely driven by statistical power, but also by considerations such as the difficulty of identifying and enrolling patients suitable for the trial and the burdens placed on patients who decide to enroll—tend to make an evidence hurdle higher. A cost-effectiveness criterion will be harder to pass when the intervention is very expensive.

Recent deliberations of the MCAC highlight differences between a purely evidence-based approach to the evaluation of medical interventions and one based on cost-effectiveness. The MCAC concluded that there was adequate evidence of effectiveness for implantable cardioverter-defibrillators; left ventricular assist devices; and verteporfin, a drug used to treat age-related macular degeneration. However, there were substantial questions about the appropriate role of each of these technologies. Published studies showed that cardioverter-defibrillators represented relatively good values for some but not all potential candidates for treatment. The cost-effectiveness of verteporfin and the left ventricular assist device had never been studied, and there was much doubt about whether the benefits were worth the high costs. The MCAC had no mechanism by which it could consider costs or value in making its coverage recommendations.

Is It Time To Rethink Evidence Evaluation As The Basis For Coverage Determinations?

Physicians, hospitals, and health plans find, as does Medicare, that evidence-based processes are not fully adequate for designing care or reimbursement policies. The idea that health plans should only pay for care that is of known effectiveness is no longer controversial. But health care providers and plans increasingly question whether medical innovations that provide genuine but modest benefits at high cost should be adopted.

Cost-effectiveness analysis has long been the preferred method to explicitly address value in medical care, yet it is not a common feature of formal coverage decision making by private U.S. health plans. My colleagues and I conducted a survey of medical directors of 228 managed care plans nationwide in 2001, representing 119 million covered lives. This survey revealed that 90 percent of the plans consider costs in some form when evaluating new interventions.²³ However, only 40 percent use formal cost-effectiveness analysis (Exhibit 1). Effectiveness appears to trump cost as an influence on coverage decisions: According to medical directors, 93 percent of all plans and 98 percent of large plans will cover a more effective intervention, even if it is more costly. If a new intervention is more expensive but no more effective than an existing one, only 16 percent will cover it, while only 8 percent will cover a less costly new intervention if it is also less effective (Exhibit 2).

Concerns about the interpretation of existing insurance contracts and about potential litigation may have discouraged the explicit use of cost-effectiveness analysis in coverage decisions.²⁴ In addition, providers and health plans may have doubts about the soundness of cost-effectiveness methods. Notwithstanding widely cited standards for the conduct of cost-effectiveness studies, questions remain about technical aspects of the methods and the ways they should be implemented.²⁵ The medical profession is not nearly as familiar with cost-effectiveness analysis as with clinical trials, and to many nonspecialists, cost-effectiveness anal-

EXHIBIT 1

How Health Plans Take Cost Into Consideration When Evaluating New Interventions

| | Formal CE analysis (%) | Selectively apply preauthorization (%) | Establish explicit coverage policies (%) | Require less costly interventions first (%) | Consider cost in any of these ways (%) |
|-------------|------------------------|--|--|---|--|
| Small plans | 36 | 48 | 53 | 62 | 92 |
| Large plans | 44 | 51 | 56 | 55 | 88 |
| All plans | 40 | 49 | 54 | 58 | 90 |

SOURCE: Adapted from L.A. Bergthold et al., "Using Evidence and Cost in Managed Care Decision-Making" (Stanford, Calif.: Center for Health Policy/Center for Primary Care and Outcomes Research, Stanford University, 2002), available as a supplemental document online at content.healthaffairs.org/cgi/content/full/hlthaff.w4.284v1/DC2.

NOTE: CE is cost-effectiveness.

EXHIBIT 2
Likelihood That Plan Will Cover A New Intervention Compared With A Standard Intervention

| | Equal effectiveness for equal cost (%) | Equal effectiveness for greater cost (%) | Less effectiveness for equal cost (%) | Less effectiveness for less cost (%) | Greater effectiveness for equal cost (%) | Greater effectiveness for greater cost (%) |
|-------------|---|---|--|---|---|---|
| Small plans | 92 | 10 | 2 | 3 | 99 | 87 |
| Large plans | 96 | 21*** | 4 | 13** | 99 | 98 |
| All plans | 94 | 16 | 3 | 8 | 99 | 93 |

SOURCE: Adapted from L.A. Bergthold et al., "Using Evidence and Cost in Managed Care Decision-Making" (Stanford, Calif.: Center for Health Policy/Center for Primary Care and Outcomes Research, Stanford University, 2002), available as a supplemental document online at content.healthaffairs.org/cgi/content/full/hlthaff.w4.284v1/DC2.

** $p < .05$ *** $p < .01$

ysis is neither transparent nor easily understood. This strikes specialists as ironic, since the technique can highlight otherwise implicit assumptions and make it easier to appreciate their implications.

Another reason for its limited adoption is the difficulty in conveying the magnitude and implications of uncertainty in the findings. There are numerous sources of uncertainty in such analyses: sample variability of outcomes observed in clinical trials; uncertainty about health events occurring after the end of a trial; uncertainty about nearly every component of costs; and uncertainty about the structure of models used in the analyses. Several techniques are available to measure and present these sources of uncertainty, including sensitivity analysis and probabilistic cost-effectiveness analysis.²⁶ The methodology and presentation of uncertainty in cost-effectiveness analysis has grown more sophisticated in recent years, but there is not a consensus about how the uncertainty should influence decision making. For all of these reasons, cost-effectiveness analysis is not poised to replace evidence evaluation in the near future.

However, it is likely that cost-effectiveness analysis will complement evidence evaluation. Although they report limited use of formal cost-effectiveness analysis and are not sure how to use it, many medical directors believe that it can and should play a greater role.²⁷ They believe that evidence evaluation should remain an important component of the decision to cover a medical product or service. However, because it does not incorporate cost considerations and because it is an imperfect proxy for cost-effectiveness, it is no longer an adequate basis for coverage decision making.²⁸ They could build upon current approaches while incorporating costs simply by assessing the cost-effectiveness of interventions that pass an evidence criterion but whose value is in question, and they could use cost-effectiveness analysis to help decide what to do when there is suggestive but not compelling evidence of effectiveness.

What action should plans take when they conclude that an intervention is not cost-effective? They could deny coverage entirely in limited circumstances, such

as a procedure that costs more and is clearly less effective (in the language of cost-effectiveness analysis, strictly dominated) than an alternative. It would be harder to deny coverage for a unique treatment for a life-threatening disease solely on the basis of poor cost-effectiveness.

There may be a broader scope for application of cost-effectiveness analysis in other aspects of benefit design. Several years ago, Mark Pauly and Philip Held argued that future cost savings from some interventions approached or even exceeded their immediate costs. For example, pneumococcal vaccine in a high-risk patient costs less than the resulting decline in spending for future care of pneumonia. A health plan could improve health outcomes and lower overall health spending by waiving the copayment—or even providing a subsidy—to ensure that such patients received the vaccine.²⁹ Tiered copayments for prescription drugs became common after Pauly and Held's paper appeared, and today it seems obvious that a similar copayment design could be applied to other medical interventions. In typical three-tier copayment arrangements, small copayments are required for the generic drugs; high copayments for brand-name, nonpreferred drugs; and intermediate copayments for brand-name, preferred drugs. The tiers, which are primarily based on drug acquisition costs, shift use toward lower-cost drugs.³⁰

A drawback of tiered copayment, however, is that the low-cost drugs it promotes are not necessarily high-value drugs; sometimes the most cost-effective drug is in the second tier, not the first, despite a higher acquisition cost. Mark Fendrick and colleagues have argued that health plans should set the copayment level (which could vary from one patient to another based on clinical characteristics) based on the benefit the intervention provides, not solely on its cost.³¹ Copayments for other forms of health care might also be adjusted for benefits. Procedures that are effective but not cost-effective in any identifiable patient population might be subject to high copayments or fixed, high coinsurance rates (percentage payments rather than fixed-dollar amounts), with no individual adjustments.

Any approach that requires different copayments for different interventions or for different patients may seem too complex to administer and understand today. Not long ago, tiered copayments for medications were criticized on the same grounds, yet they are ubiquitous today. As spending continues to rise, employers, consumers, and health plans will become more willing to explore alternatives to traditional health insurance.

In the absence of a return to heavily managed care or the adoption of novel approaches to coverage policy, commercial health plans are expected to continue to shift more costs to the insured, giving individuals a larger stake in the costs of the care they use. According to the 2000 and 2003 Henry J. Kaiser Family Foundation/Health Research and Educational Trust Surveys of Employer-Sponsored Health Benefits, out-of-pocket spending grew dramatically during the study years. Deductibles in preferred provider organization (PPO) plans grew by 57 percent (pre-

“As they bear more of the cost of care out of pocket, consumers may become the main audience for information about value.”

.....

ferred provider) and 65 percent (nonpreferred provider) in that time frame, while prescription drug copayments grew by 46 percent (preferred drugs) and 71 percent (nonpreferred drugs). The limits on out-of-pocket payments are also rising, and fixed coinsurance rates for hospitalizations and other costly forms of care will give patients further incentives to take costs into account. Cost sharing is an even more prominent feature of “self-directed” or “consumer-directed” health plans. As they bear more of the cost of care out of pocket in both absolute and relative terms, consumers may become the main audience for the information about value that cost-effectiveness analysis can provide.

Commercial health plans incorporate value considerations into benefit design in different ways as they compete for subscribers. One person could choose a plan that consistently applies cost-effectiveness criteria to coverage and other aspects of benefit design, providing a high-value, low-cost package. Another could choose a plan that either uses less restrictive cost-effectiveness criteria (that is, reimburses care with a less favorable cost-effectiveness ratio) and is broader in its coverage, or perhaps uses more traditional benefit design, charging higher premiums and using cost sharing more heavily. Although state and federal regulation limit the scope of plan variation, the market might help sort out which approaches have the greatest appeal to consumers.

Most government health insurance programs—Medicare Advantage plans are an important exception—do not compete in the same way that commercial health insurance plans do. Many, like Medicaid, serve a diverse and often vulnerable population, so extensive cost sharing is not feasible. Cost-effectiveness analysis will likely have a different role in these settings. State Medicaid programs use mechanisms such as capitation, low reimbursement rates, and restrictive coverage to control costs. Cost-effectiveness calculations undoubtedly enter into some of their benefit decisions, as happened so explicitly in Oregon. Medicaid programs will be more likely to pursue cost-effectiveness analysis as a basis for new approaches to benefit design if they face severe financial stresses. They might then conclude that their current formulas cannot control costs and yield acceptable health outcomes, and that formal cost-effectiveness analysis would provide useful guidance, particularly in setting limits on covered products and services.

Medicare is also a government program whose features are determined by legislation and regulatory interpretation. The use of cost-effectiveness in benefit design will be determined, therefore, by what is politically acceptable. Medicare differs from Medicaid in a very important respect: Because they are a large and politically powerful constituency, Medicare beneficiaries have a powerful voice in deliberations over any major change. The opposition of some Medicare beneficia-

ries, as well as several other influential constituencies, stymied Medicare officials' repeated attempts to introduce cost-effectiveness or even explicit consideration of cost in their coverage decision making. Past failures do not mean that every future effort of this kind will fail, though. The Medicare Prescription Drug, Improvement, and Modernization Act (MMA) of 2003 will have diverse effects, many of them unknowable, but among the certainties is its commitment of half a trillion dollars in additional federal funds to Medicare over the next ten years. By the end of that period, large numbers of baby boomers will have become eligible for Medicare. As the repercussions of this demographic phenomenon are felt and it becomes untenable to claim that costs can or should be ignored, the terms of debate about reform to Medicare benefit design may shift dramatically.

COST-EFFECTIVENESS ANALYSIS is a decades-old technique that has been studied more than it has been applied. Although it is not without flaws, it was never widely applied to U.S. coverage decisions because there was neither a consensus about how it should be used nor strong enough incentives to adopt it. The erosion of commercial health insurance and the growing burden of public health insurance programs may transform it from an academic curiosity to an essential tool for health care decision making.

.....
This paper is based in part on work supported by Robert Wood Johnson Foundation Grant no. 039396, National Institute on Aging Grant no. AG 17253, and the Laughlin Endowment. The author is grateful to Linda Bergthold and Jay Bhattacharya for their helpful comments.

NOTES

1. See A.M. Garber, "Evidence-Based Coverage Policy," *Health Affairs* 20, no. 5 (2001): 62–82.
2. Classic descriptions are in K.J. Arrow, "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review* 53, no. 5 (1963): 941–973; and M.V. Pauly, "The Economics of Moral Hazard: Comment," *American Economic Review* 58, no. 3 (1968): 531–537.
3. See B.A. Weisbrod, "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care and Cost Containment," *Journal of Economic Literature* 29, no. 2 (1991): 523–532; and S.T. Burner and D.R. Waldo, "National Health Expenditure Projections, 1994–2005," *Health Care Financing Review* 16, no. 4 (1995): 221–242.
4. See I. Ehrlich and G.S. Becker, "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy* 80, no. 4 (1972): 623–648.
5. See discussions in L.A. Bergthold, "Medical Necessity: Do We Need It?" *Health Affairs* 14, no. 4 (1995): 180–190; S.J. Singer and L.A. Bergthold, "Prospects for Improved Decision Making about Medical Necessity," *Health Affairs* 20, no. 1 (2001): 200–206; and L.A. Bergthold et al., "Using Evidence and Cost in Managed Care Decision-Making" (Stanford, Calif.: Center for Health Policy/Center for Primary Care and Outcomes Research, Stanford University, 2002), available online at content.healthaffairs.org/cgi/content/full/hlthaffw4.284v1/DC2.
6. See D.M. Eddy, "Benefit Language: Criteria That Will Improve Quality While Reducing Costs," *Journal of the American Medical Association* 275, no. 8 (1996): 650–657; and D.M. Eddy, "Investigational Treatments: How Strict Should We Be?" *Journal of the American Medical Association* 278, no. 3 (1997): 179–185.
7. The processes Blue Cross Blue Shield uses are described in S. Gleeson, "Blue Cross and Blue Shield Association Initiatives in Technology Assessment," in *Adopting New Medical Technology*, ed. A.C. Gelijs and H.V. Dawkins (Washington: National Academies Press, 1994). The Medicare Coverage Advisory Committee (MCAC) is described in Health Care Financing Administration, "Procedures for Making Coverage Deci-

sions,” *Federal Register* 64, no. 80 (1999): 22619–22625. There are undoubtedly many reasons for the acceptance of evidence-based processes. Among them are the recognition that there are widespread variations in practice patterns that cannot be explained by patient characteristics alone and that clinical trials and other high-quality clinical studies are now common, so it seems more feasible than in the past to meet an evidence standard.

8. See U.S. Preventive Services Task Force, *Guide to Clinical Preventive Services*, 2d ed. (Baltimore: Williams and Wilkins, 1996); and Canadian Task Force on the Periodic Health Examination, “The Periodic Health Examination: Canadian Task Force on the Periodic Health Examination,” *Canadian Medical Association Journal* 121, no. 9 (1979): 1193–1254.
9. See J. Concato, N. Shah, and R.I. Horwitz, “Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs,” *New England Journal of Medicine* 342, no. 25 (2000): 1887–1892; K. Benson and A.J. Hartz, “A Comparison of Observational Studies and Randomized, Controlled Trials,” *New England Journal of Medicine* 342, no. 25 (2000): 1878–1886; and M.A. Hlatky et al., “Comparison of Predictions Based on Observational Data with the Results of Randomized Controlled Clinical Trials of Coronary Artery Bypass Surgery,” *Journal of the American College of Cardiology* 11, no. 2 (1988): 237–245.
10. Cost estimates are from J.A. DiMasi, R.W. Hansen, and H.G. Grabowski, “The Price of Innovation: New Estimates of Drug Development Costs,” *Journal of Health Economics* 22, no. 3 (2003): 151–185.
11. See D. Moher, C.S. Dulberg, and G.A. Wells, “Statistical Power, Sample Size, and Their Reporting in Randomized Controlled Trials,” *Journal of the American Medical Association* 272, no. 2 (1994): 122–124.
12. Increasing the number of patients enrolled is only one of the mechanisms to ensure a large enough number of observed events, which drive the power of the trial. For example, investigators can make great efforts to improve the completeness of reporting of all health events, and they can work to minimize the number of people who drop out of a trial or are lost to follow-up. Investigators also try to enroll only those patients who are likely to adhere to all aspects of demanding protocols for participation in the trial, improving the chances that the treatment will be used properly and its effects observed. These and other aspects of trial design that tend to increase statistical power, while increasing the credibility of study results, are labor-intensive.
13. For many devices, the evidence barrier (both to approval and to the entry of new competitors) has been much lower than for pharmaceuticals, so large, well-designed randomized trials are more common for drugs than for devices.
14. Evidence from the past ten to fifteen years suggests that team care—or “disease management”—is often the most effective approach to the management of chronic diseases. Chronic disease management typically requires selecting a portfolio of diagnostic, monitoring, and treatment strategies, tailored to the individual patient, rather than simply dispensing a medication and obtaining occasional laboratory tests. Although some programs use proprietary software or are provided by dedicated disease management companies, the key features of disease management are matters of public knowledge. Because the benefits of research in these strategies are difficult for any individual firm to capture, randomized trials of disease management are less common than trials of drugs and medical devices. Furthermore, reimbursement for disease management was slow to develop, especially among fee-for-service insurers. According to the McKinsey Global Health Care Productivity study, disease management for diabetes reduced costs of care and improved outcomes. Such programs were adopted earlier in the United Kingdom than in the United States; slower U.S. adoption seemed to reflect the absence of reimbursement for components of diabetes team care. See M.N. Baily and A.M. Garber, “Health Care Productivity,” *Brookings Papers on Economic Activity: Microeconomics* (1997): 143–202.
15. See M.V. Pauly, “Should We Be Worried about High Real Medical Spending Growth in the United States?” *Health Affairs*, 8 January 2003, content.healthaffairs.org/cgi/content/abstract/hlthaff.w3.15 (7 April 2004); and Burner and Waldo, “National Health Expenditure Projections, 1994–2005.”
16. See M.C. Weinstein and W.B. Stason, “Foundations of Cost-Effectiveness Analysis for Health and Medical Practices,” *New England Journal of Medicine* 296, no. 13 (1977): 716–721; and D.M. Eddy, “Cost-Effectiveness Analysis: A Conversation with My Father,” *Journal of the American Medical Association* 267, no. 12 (1992): 1669–1675.
17. See A.M. Garber et al., “Theoretical Foundations of Cost-Effectiveness Analysis,” in *Cost-Effectiveness in Health and Medicine*, ed. M.R. Gold et al. (New York: Oxford University Press, 1996); and C.E. Phelps and A.I. Mushlin, “On the (Near) Equivalence of Cost Effectiveness and Cost Benefit Analysis,” *International Journal of Technology Assessment in Health Care* 7, no. 1 (1991): 12–21.
18. A.M. Garber and C.E. Phelps, “Economic Foundations of Cost-Effectiveness Analysis,” *Journal of Health Eco-*

- nomics 16, no. 1 (1997): 1–31.
19. S. Birch and A. Gafni, “Cost-Effectiveness Ratios: In a League of Their Own,” *Health Policy* 28, no. 2 (1994): 133–141; and M. Drummond, J. Mason, and G. Torrance, “Cost-Effectiveness League Tables: Think of the Fans,” *Health Policy* 31, no. 3 (1995): 231–238.
 20. A comprehensive listing of cost-effectiveness ratios, with comments on characteristics of the studies used to generate the numbers, can be found at the Harvard Center for Risk Analysis Web site, www.hsph.harvard.edu/cearegistry/ (6 May 2004).
 21. See J.A. Kitzhaber, “Prioritising Health Services in an Era of Limits: The Oregon Experience,” *British Medical Journal* 307, no. 6900 (1993): 373–377; D.M. Eddy, “Oregon’s Methods: Did Cost-Effectiveness Analysis Fail?” *Journal of the American Medical Association* 266, no. 15 (1991): 2135–2141; and T.O. Tengs et al., “Oregon’s Medicaid Ranking and Cost-Effectiveness: Is There Any Relationship?” *Medical Decision Making* 16, no. 2 (1996): 99–107.
 22. The definition of “cost” often determines whether an intervention is considered “expensive,” and to whom. Most pharmaceutical products have prices that are very high compared with the marginal cost of production. If “cost” in the cost-effectiveness analysis refers to the retail price, such a drug will often pass an evidence criterion more readily than a cost-effectiveness criterion. If the cost of production is high relative to the price, as would often be the case for a complex surgical procedure, it may be relatively difficult to pass an evidence criterion, since there would be so little return to an investment in studies demonstrating effectiveness. This would even be true for a procedure that was highly cost-effective.
 23. The survey was mailed to the medical directors of 346 eligible managed care plans in 49 states and the District of Columbia; the 66 percent of plans that responded were responsible for the care of 77 percent of the members of the 346 plans in the sample. The survey instrument was a closed-ended mail questionnaire consisting of forty-two questions divided into seven topic areas, including evaluation of clinical effectiveness and evaluation of cost and cost-effectiveness. Details of the survey and its methods are in Bergthold et al., “Using Evidence and Cost.”
 24. D.M. Eddy, “The Use of Evidence and Cost-Effectiveness by the Courts: How Can It Help Improve Health Care?” *Journal of Health Politics, Policy and Law* 26, no. 2 (2001): 387–408; and P.D. Jacobson and M.L. Kanna, “Cost-Effectiveness Analysis in the Courts: Recent Trends and Future Prospects,” *Journal of Health Politics, Policy and Law* 25, no. 2 (2001): 291–326.
 25. The recommendations of the federally sponsored Panel on Cost-Effectiveness in Health and Medicine appear in Gold et al., eds., *Cost-Effectiveness in Health and Medicine*.
 26. Leading studies of methods for valuing uncertainty appear in J. Mullahy and W.G. Manning, “Statistical Issues in Cost-Effectiveness Analysis,” in *Valuing Health Care: Costs, Benefits, and Effectiveness of Pharmaceuticals and Other Medical Technologies*, ed. F. Sloan (New York: Cambridge University Press, 1994); B.J. O’Brien et al., “In Search of Power and Significance: Issues in the Design and Analysis of Stochastic Cost-Effectiveness Studies in Health Care,” *Medical Care* 32, no. 2 (1994): 150–163; P. Wakker and M.P. Klaassen, “Confidence Intervals for Cost/Effectiveness Ratios,” *Health Economics* 4, no. 5 (1995): 373–381; A. Briggs and M. Sculpher, “Sensitivity Analysis in Economic Evaluation: A Review of Published Studies,” *Health Economics* 4, no. 5 (1995): 355–371; and A. Briggs, M. Sculpher, and M. Buxton, “Uncertainty in the Economic Evaluation of Health Care Technologies: The Role of Sensitivity Analysis,” *Health Economics* 3, no. 2 (1994): 95–104.
 27. Bergthold et al., “Using Evidence and Cost.”
 28. See N. Daniels and J.E. Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?* (New York: Oxford University Press, 2002).
 29. M.V. Pauly and P.J. Held, “Benign Moral Hazard and the Cost-Effectiveness Analysis of Insurance Coverage,” *Journal of Health Economics* 9, no. 4 (1990): 447–461.
 30. B. Motheral and K. Fairman, “Effect of a Three-Tier Prescription Copay on Pharmaceutical and Other Medical Utilization,” *Medical Care* 39, no. 12 (2001): 1293–1304.
 31. A.M. Fendrick et al., “A Benefit-Based Copay for Prescription Drugs: Patient Contribution Based on Total Benefits, Not Drug Acquisition Cost,” *American Journal of Managed Care* 7, no. 9 (2001): 861–867.