

# The Promise and the Problem

The good news is that modern medicine works miracles. The bad news is that it breaks banks—public and private. The benefits from improved health care exceed by trillions of dollars its admittedly large and growing cost.<sup>1</sup> Before the late 1960s and the development of durable artificial hips, for example, it was impossible to enable people crippled by painful, arthritic hips to walk normally again. Before the invention of computed tomography (CT) scanners or magnetic resonance imaging (MRI) it was often impossible—short of invasive, painful, and costly exploratory surgery—to pinpoint the exact location of many deep tumors or abscesses. Such procedures generate total costs far higher than the more primitive techniques they replaced but produce enormous gains for many patients. Few Americans would willingly trade today's health care, costly though it is, for the less expensive but less effective treatments of the past.

Still, the cost of these advances is staggering. Real U.S. medical expenditures have increased sevenfold since 1965, when Congress passed Medicare and Medicaid. Outlays will reach \$11,046 per person and 18.7 percent of gross domestic product by 2014, according to official projections.<sup>2</sup> If the gap between growth of health care spending and income persists, the former would claim half of all increases in income by 2022 and all of it by 2051. Total health care spending would claim about 28 percent of total U.S. production by 2030 and 35 percent by 2040.<sup>3</sup>

Such rapid growth in spending is not unique—for example, real computer purchases have risen 188-fold since 1978.<sup>4</sup> Although far faster than the increase in health care spending, the increase in computer spending evokes no fretful hand wringing about spiraling computer expenditures. No politician feels driven to orate on the need to contain computer costs. To understand the difference, one need look no further than health insurance. Computer buyers pay directly for each additional machine they buy. Health care consumers do not: on average, patients pay out-of-pocket for only 3 percent of hospital care and 11 percent of physician services.<sup>5</sup> And during serious illnesses, when out-of-pocket spending exceeds defined limits, many insurance plans now pay all costs of care. To be sure, when health costs rise, the cost of health insurance—private and public—goes up. But the prospect of such charges has little influence on the amount of care that well-insured patients seek when ill or that their caregivers are disposed to render.

Insurance protects people from financial ruination by health care costs. Indeed, it is the lack of such coverage by some 45 million people that troubles elected officials and everyone else, even if agreement on how to extend coverage has proven elusive. In addition to—indeed, precisely because of—such financial protection, insurance produces two side effects. First, by shielding patients from all or most of the cost of care, insurance encourages patients to demand all care, however small the benefit and however high the cost of producing it. Sometimes such care—the test that provides additional information of little value, the medicine that is trivially better than a much less costly alternative, the surgery that is expected to produce only small improvements—generates benefits that are, in some sense, smaller than the costs.<sup>6</sup> And as medical technology improves and total benefits from advancing technology increase, outlays on care that does not deliver value for money also will tend to increase. Second, health insurance also relieves biomedical investigators of any need to worry about the cost of new and better treatments.

Victor Fuchs and Alan Garber vividly illustrate the working of these two unintended side effects of insurance:

Imagine how the market for automobiles would have developed if a third party had provided automobile insurance that paid 80 percent of the cost of new cars. . . . Such insurance would influence both the number and types of cars people bought. People would replace cars more often, and they would buy higher quality cars

than they do in today's automobile market. A Lincoln or a Mercedes would cost buyers little more than a Chevrolet or a Honda, and sales of luxury automobiles would rise. Auto manufacturers would focus their product development on quality enhancements, such as big engines and luxurious interiors, rather than on cost-reducing manufacturing changes. . . . The quality-adjusted price of automobiles might fall, but since only high-quality cars would be sold, the average price of a car would rise. . . . The well-insured might welcome the steady improvements in the quality of luxury cars, but many would be better off with simpler automobiles and higher take-home pay.<sup>7</sup>

The enormous benefits and costs of prospective medical advances magnify the importance of reaching three goals. The first is to ensure that all Americans enjoy financial access to modern health care. That means extending health insurance to the currently uninsured and upgrading coverage for those who now are inadequately insured. Achieving this goal will doubtless add to health care spending. The second goal is to ensure that all patients receive care that is worth what it costs. The failure of the U.S. health care system to meet this standard, even for the well insured, is increasingly well documented. The third goal is to eliminate care that is not worth what it costs.

Achievement of universal coverage is likely to be a necessary precursor to eliminating care that is not worth what it costs. The uninsured now receive substantial amounts of free or subsidized care that is financed by extra charges levied on the insured. Such cross-subsidies provide the modicum of protection for the uninsured without which the status of being uninsured would be intolerable. Yet aggressive efforts to control costs for the insured inevitably constrict the capacity of health care providers to sustain such "unreimbursed" care. Effective cost control would also require control mechanisms that do not currently exist in the United States.

The practical questions concern how to curtail such care. Public policy could limit demand for care—for example, by withdrawing tax concessions that lower the cost of health insurance or by raising the price of care through higher premiums or increased charges at the time of use. The legislation enacted in 2003 to encourage high-deductible insurance is an example of such an approach. However, demand restriction suffers from a serious shortcoming. Any insurance that effectively protects people

from financial catastrophe must pay for essentially all costs beyond a limit. Most health costs are incurred by a small proportion of the population whose expenses greatly exceed plausible limits on out-of-pocket spending. Demand-limiting strategies can deter spending up to a point but can exercise no influence at the margin on care provided once costs exceed such limits. For that reason, demand limitations are likely to have only a modest impact on the development and deployment of new medical technologies.

The second general approach, to curb the advance of medical technology—for example, by curtailing public support of basic science—would be neither desirable nor effective. It is not desirable because, as noted, the average benefits of advancing medical technology are likely in the future, as they have been in the past, to greatly exceed their costs. It is not likely to be effective because the United States is not the only sponsor of medical research. Other nations would be ready and are increasingly able to entice top investigators with research support if the United States cuts off such funding. Therefore, curtailing biomedical research in the United States would sacrifice scientific leadership without achieving the intended reduction in cost-increasing health technology.

The third approach to controlling growth in health care spending is to try to limit the provision of care that is worth less than it costs people who are well insured—that is, to ration care. Other nations commonly use budgetary limits or restrict the supply of key personnel or equipment so that health care providers must limit care. The current maze of ways Americans pay for health care is incapable of enforcing such limits. To create such controls would precipitate passionate debate centered on deep ideological divisions over the proper scope for collective authority and individual rights.

Thus the steps necessary to ration health care may prove more objectionable than the cost of paying for it. But as the health care menu lengthens, the cost of low-benefit care that well-insured patients will demand under current payment arrangements is likely to become so burdensome that serious examination of health care rationing and what it would take to implement it will prove inescapable. The outcome of such a debate cannot be prejudged. But to discuss the options intelligently, Americans need to understand not just the size of the economic challenge but also the choices that health care rationing will entail. This book is an attempt to provide readers with such information.

## Some History

Rapid growth of health care expenditures has led successive presidents to propose limits on hospital outlays or other medical spending. The administrations of presidents Nixon, Ford, Carter, Reagan, and Clinton proposed to limit growth of health care spending through a number of means: price controls (Nixon), hospital revenue controls (Nixon and Carter), limits on Medicare and Medicaid spending, and increased competition (Reagan and George W. Bush), or sweeping reform (Clinton). On a parallel track, Congress tried to slow growth of medical expenditures by requiring states to discourage duplication of medical facilities by having care providers obtain a certificate of need from the local health system agencies before making a capital expenditure greater than \$100,000 or \$150,000. State governments introduced mandatory controls on hospital charges, third-party payment rates, or total hospital revenue. The state programs all proved unsustainable and lapsed or languished because none succeeded in limiting outlays more than briefly.

The two most enduring efforts to control costs have been Medicare's shift from cost-based reimbursement to prospective payment systems and private sector adoption of managed care. Until 1983 Medicare paid hospitals based on costs incurred. In that year it began to pay hospitals fixed amounts set prospectively for most in-patient services and based on patients' diagnoses at time of admission. These fees have been adjusted annually based on changes in the cost of hospitals' inputs, hospital productivity, and other considerations. Initial prices were quite generous, but Medicare gradually used this system to slow spending growth. Medicare later established fee schedules for physician services. Still later it extended prospective payment to hospital outpatient services, home health providers, and skilled nursing facilities.<sup>8</sup>

The second cost control effort played out in the private sector. In the 1990s private payers began to try to control costs through private regulatory devices grouped under the term *managed care*. This term includes the cost control methods of health maintenance organizations that provide all or most care through salaried or contractual staff. It encompasses preferred provider organizations that encourage enrollees to use designated providers who offer discounts. It also covers the efforts by insurers or plan managers to limit fees, screen patients for particular services, and negotiate discounts from suppliers, physicians, and other

health professionals. At about the time that these practices became widespread, growth of health care spending decelerated. Whether this slowdown resulted from or simply coincided with managed care is unclear. That the slowdown was short lived is not. People objected to the aggressive cost control methods used by managed care companies. Personal testimonies, news reports, and movies made managed care an object of fear and loathing for many people who objected to the denial of care by business executives and even physicians who stood to gain financially from their parsimony. The plans scaled back efforts to control costs.

Rapid growth of per capita health care spending—which has outpaced income growth by an average of 2.5 percent annually since 1960—resumed in 2000. This excess growth seems likely to persist until and unless the U.S. public is prepared to tolerate effective cost limits, private or public. To be sure, spending growth has slowed episodically in the past and may do so in the future. The reduction of wasteful practices could achieve sizable, but one-time, savings. A slowdown in the advance of medical technology and the attendant flow of new therapeutic and diagnostic procedures would also slow the growth of health care spending. But such a slowdown seems highly improbable in light of the recent breakthroughs in molecular and cellular biology and in information technology. More importantly, it would be a misfortune because it would mean that humans were deprived of the life-extending, pain-reducing, and function-enhancing contributions that have defined medical advance.

## Rationing

What this all means is that a *sustained* slowdown in the growth of health care spending will require rationing—the denial of some beneficial care to some people who have the financial means to pay for it. Such rationing should not be confused with the absence of health insurance for millions of Americans. Many observers characterize this situation as “rationing by price.” But the term *rationing* ordinarily does not refer to the inability of some people in a market economy to afford particular commodities. Customarily, it describes the situation in which people who can afford a commodity are unable to buy it because of scarcity, which results because some nonmarket allocation system—ration coupons or queues, for example—limits demand to available supply.

No other nation spends nearly as much on health care as does the United States. Per capita health care spending in the United States, at

\$5,267 in 2002, was more than twice the \$2,049 average of other members of the Organization for Economic Cooperation and Development (OECD) and more than half again as great as spending in the second highest spending nation—Switzerland (\$3,446).<sup>9</sup> These differentials are attributable to several causes. First, U.S. physicians receive particularly generous remuneration. Second, rich nations may spend a larger share of their incomes on health care than do poorer nations, and the United States is richer than most other OECD nations.<sup>10</sup> Third, the United States tends to have more medical equipment and higher rates of surgery than do most other nations.

Other wealthy nations limit health care spending in various ways. The highly decentralized U.S. payment system is unique in its lack of effective levers for limiting health care spending. The theme of this book is that the United States will be forced by sharply rising public and private health care spending to consider the adoption of such limits. Even if the total benefits from advancing medical technology far outweigh the total costs, the amount spent on care that provides only marginal benefits is likely to grow even faster (see chapter 6). For this reason, the stakes in controlling health care spending will increase even as the benefits from advancing medical technology grow. Limits, whatever their particular form, will require that some sick people be denied some care that is somewhat beneficial but worth less than it costs.

To understand whether the rationing “cure” is less painful than the “disease” of sharply higher taxes and private health care spending, one must understand what sorts of trade-offs rationing requires. U.S. experience offers scant guidance because the well-insured now enjoy financial access to essentially all beneficial care. Nor do the uninsured offer a good indication of what rationing would entail, as whatever care they receive comes from hospitals and physicians whose practice patterns are shaped by the financial incentives resulting from the majority, who are well insured.

To secure insight on the implications of health care rationing, one must look at a system in which resources are limited for all. The British health care system provides such a perspective. The United States is unlikely to adopt the particular institutional arrangements that the British have used to control health care spending or to ration care as severely. Other developed nations employ different methods for controlling spending, and few have rationed care as stringently as have the British. For that reason, the British experience provides a clearer view of the kinds of trade-offs that

rationing entails. The final chapter describes various ways in which rationing might be implemented in the United States and the results likely in the U.S. context. Before then, however, it is essential to see what rationing would entail.

### Lessons from Britain

Important similarities between Britain and the United States suggest that British reactions to resource limits will provide some guidance to probable U.S. reactions to such limits. Each country's medical journals are commonly read in the other. The clinical and scientific standards of both nations are similar. Some physicians from each country spend time in the other as students, teachers, and researchers.

Despite these similarities, British and American societies are not identical. Britain has rigidly controlled medical expenditures for decades.<sup>11</sup> Per capita hospital expenditures are now about 45 percent of those in the United States.<sup>12</sup> For decades such limits have shaped British medical practice and patient attitudes, forcing the British to decide what medical services *not* to provide, a challenge that the United States so far has declined to confront. The organization of health care, the political system, and the relative importance of class differ between the two nations. British patients are less demanding of their health care providers and less litigious than Americans, although these dissimilarities may be narrowing. For these reasons, British behavior is not an exact model for the choices Americans would make if the United States sought to severely curb medical expenditures.

Chapter 2 describes the British health care system for American readers. To measure the practical results of budget limits, chapters 3 through 5 compare the provision of representative health procedures in the United States and Britain.<sup>13</sup> This book is based on the assumption that U.S. service levels provide a benchmark for treatment levels if well-insured patients receive all care expected to generate net medical benefits. To be sure, Medicare's prospective payment system, private managed care, and capitated health plans can discourage the provision of some beneficial care. Medicare's prospective payment system means that hospitals do not receive larger payments if they provide additional services (other than extra payments for unusually costly cases). Some health care organizations use financial incentives to discourage what they regard as excessive care. Others require prior approval for some treatments as a condition for



payment. But the large majority of Americans receive care under plans that have few tools to control spending. The system of cost-based reimbursement encourages the provision to most patients of all care that promises to yield benefits, regardless of cost.

In some instances, the U.S. system provides more than all beneficial care. Huge geographic variations in the provision of many forms of health care persist that differences in rates of illness cannot explain. Studies indicate that outcomes in regions of particularly heavy use are no better than elsewhere.<sup>14</sup> Such evidence strongly suggests that too much care is provided in some places. The fear of being sued may cause some physicians to practice defensive medicine, that is, to provide care designed more to minimize the risk of being sued than to improve patient outcomes. Other physicians may perform surgery they consider of marginal value or recommend unnecessary tests in order to boost their own incomes. The extent or even existence of these practices remains a matter of controversy.<sup>15</sup> On balance, however, care in the United States for most well-insured patients still remains close to what would be provided if cost were no object and benefit to patients were the sole concern.

Given this assumption, the British system is deemed to provide “full care” if its service levels are similar to those in the United States. The difference is a measure of rationing. The practical question is whether those differences make medical and social sense. Chapter 6 lays out an analytical framework for thinking about these questions and presents estimates of the degree to which differences in the provision of the procedures examined in chapters 3 through 5 account for the large difference between U.S. and British health care expenditures.

### **What to Look For**

Chapters 3, 4, and 5 reveal gaps of widely varying sizes between the United States and Britain in the availability of several medical services. These chapters also explore possible medical justifications or other explanations for these differences. The data are inadequate to test in a statistically rigorous way precisely why the British limit various forms of care to such different degrees. Readers should decide whether the reasons presented here are plausible. They should also decide whether they think that people in the United States would respond similarly or differently.

Some observed behaviors are ones that any budget-limited system would elicit. Physicians who say “no” must learn to do so in ways that

are acceptable both to themselves and to their patients. Patients unwilling to accept the consequences of resource limits are likely to seek ways to “work the system” to secure care they were initially denied. Communities may circumvent limits by donating equipment that would not otherwise be available. Interest groups may use the media to try to pressure the government to increase allocations for the treatment of certain diseases. Chapter 7 summarizes British responses to limits on health care, while chapter 8 considers how such responses would affect the operation of budget limits in the United States.

Chapters 7 and 8 also address several other complex questions. Should patients be permitted to buy medical care outside a system subject to budget limits? Can clinical freedom survive in an environment of budget limits? Should charitable gifts always be welcomed? What legal actions would arise because of effective budget limits? How should limits be structured so that physicians and patients operate within them, rather than working to defeat them?

Control of health care spending involves many technical issues. For example, hospital spending can be controlled with fixed budgets, revenue limits per patient day, or revenue limits per admission. Alternatively, demand for hospitalization can be influenced by deductibles and co-payments of various kinds. Each control mechanism creates particular incentives and distortions. The choice among them is of fundamental importance. This volume does not deal with such issues. The focus, rather, is on those decisions and trade-offs that must be made if budget limits are to be effective. Primary emphases are on behavioral adjustments that must be encouraged, institutional changes that would result, the value judgments that budget limits would require, and the coping mechanisms that they would elicit. How far the United States will—or should—venture in rationing medical care is unclear. That it will have to confront this difficult problem is beyond dispute.