THE DYNAMICS OF TECHNOLOGICAL CHANGE IN MEDICINE

by Annetine Gelijns and Nathan Rosenberg

Prologue: Within public policy circles, high-technology medical care has been viewed as a curse and a blessing because of its capacity to consume an ever-increasing share of resources and the wonders it works in the diagnosis and treatment of disease. Although opinions abound about its value, far less is understood about the processes of innovation and the multiple forces that encourage or impede technologic development. In this paper Annetine Gelijns and Nathan Rosenberg discuss these processes, underscoring the diverse influences to which they are subjected. Gelijns is an assistant professor in the Department of Surgery and School of Public Health at Columbia University. She is associate director of the Center for Health Outcomes and Innovation Research in Surgery there. Gelijns holds a doctorate in medicine and science policy from the University of Amsterdam and a master of law degree from the University of Leyden, the Netherlands. Rosenberg is a professor of economics at Stanford University. He has devoted much of his career to the study of innovation in different industrial sectors and is considered to be a preeminent authority on the economics of technological change. Gelijns and Rosenberg describe innovation as a dynamic process that is greatly influenced by a close interaction between developers and users, by changes in financing and delivery modes, and by shifting patterns of medical specialization. The processes of innovation involve a wide range of actors and interests, which operate under different incentives and time lines. This paper argues that it is not possible to establish the causal connections between technological change and rising medical costs without first understanding the dynamics of innovation.
The health care reform debate in the United States is focusing attention on technological change, mainly because it has been widely identified as the principal culprit in the rising costs of medical care. Authoritative analysts such as Victor Fuchs have written of a “technological imperative” driving up the cost of medical care through its influence on medical choices. More recently Joseph Newhouse has suggested that more than 50 percent of the total rise in real medical care costs (after inflation) is attributable to technological changes. This appears to contradict conventional wisdom. Outside of medicine, technological change is identified as the primary driving force behind improved productivity and economic growth. One of the most decisive effects of technological change is that it makes it possible to produce a given volume of output with a smaller volume of inputs. Why, then, when considering medicine, is technological change deemed responsible for rising costs?

Addressing this apparent anomaly, first of all, requires careful examination of the static functioning of the health care system—how it allocates and uses the resources available at a particular moment in time. In the past the operation of this system has been subject to a number of distortions that have flowed primarily from (1) the extreme information asymmetry between physician and patient that has allowed the physician to determine the demand for medical services; and (2) the fact that third-party payment has insulated patients and physicians from the financial implications of their medical care decisions. In an environment of considerable uncertainty concerning the real value of medical interventions, physicians have been able to expand the demand for medical services, an expansion that has been intensified by a concern over malpractice suits. Such “defensive medicine” has also been consistent with the financial, sociocultural, and professional incentives encouraging the broad application of technology. Recently, these distortions have led to suggestions for “fixing” the system, including more systematic copayments, higher deductibles, managed competition, and the like.

However, several dynamic issues involving the ways in which new technologies are developed, adopted, and adapted are equally critical to long run containment of rising medical costs. Surprisingly, the processes through which new medical procedures and products are generated and introduced
have received much less attention in the current health care reform debate. The primary purpose of this paper is to examine the underlying dynamics of technological innovation in medicine. In our view, it is not possible to establish causal connections between technological change and rising medical costs without first understanding these dynamics. Once these forces have been grasped more clearly, the relationship between technology and medical care costs will be more powerfully illuminated.

Some Salient Characteristics Of Medical Innovation

The popular perception of medical innovation is one in which a group of biomedical scientists has a bright new idea, which then moves in a linear progression from the laboratory to animal models, to select populations, and finally to the bedside (Exhibit 1). Although much innovation in medicine flows in this fashion, this linear conceptualization captures only part of the reality, particularly with regard to medical devices. A high percentage of new medical devices have emerged not out of biomedical research, but through transfer of technologies that were developed elsewhere: lasers, ultrasound, magnetic resonance spectroscopy, and that most general-purpose of all technologies, the computer. Both the modern intensive care unit, with its dependence upon elaborate monitoring technologies, and much medical research itself have been based upon new capabilities acquired through transfer and subsequent modification to suit the particular needs of the medical sector. Indeed, it may be much closer to the truth to say that these electronic technologies have vastly strengthened the capacity to perform “upstream” biomedical research than that they were the results of such research in the first place. Magnetic resonance imaging (MRI), for example, a technology that had its origins in basic research on the structure of the atom, later was transformed into a major diagnostic tool. This diagnostic tool, in turn, has vastly improved the ability to research various internal organs—thus highlighting the nonlinear nature of the innovation process.

Another drawback of the linear model is that it implies that one can make a neat distinction between research and development (R&D) on the one hand and adoption on the other, with all of the uncertainty inherent in
innovation attached to the former. Certainly, if one looks at the whole spectrum of R&D, the uncertainties appear to be larger the more one moves toward the basic research end of the spectrum. But it is a serious misperception to think that all important uncertainties have been ironed out by the time a new technology has finally been introduced into clinical practice.\(^3\) In fact, much uncertainty associated with a new technology can be resolved only after extensive use in practice. Thus, development does not end with the adoption of an innovation. Actual adoption constitutes only the beginning of an often prolonged process in which important redesigning takes place, exploiting the feedback of new information generated by users.\(^4\)

Development of the “pill” is a case in point. The widespread introduction of the first oral contraceptives into clinical practice during the early 1960s confirmed their high degree of effectiveness but also revealed that their use increased the risk for thromboembolic disorders. The suspicion, based on subsequent research, that estrogen might be responsible for such circulatory diseases stimulated manufacturers to reduce estrogen levels and develop low-dose pills, which led to a dramatic decline in side effects. Because the first drug in a new therapeutic class is probably never the optimal version, incremental improvements after initial adoption play an important role in pharmaceutical and biological development.

Medical devices are characterized by even higher levels of incremental change than is true for therapeutic drugs. Consider, for instance, the evolution of endoscopes. Today’s “cold-light” video-endoscope, with a computer-chip camera at its tip and with capabilities that allow its use both for diagnosis and therapy, is a world apart from its predecessor of the early 1950s. During those years, for instance, the lamp at the tip of the endoscope could cause serious burns, vision was often restricted, the quality of images was poor, therapeutic applications were essentially nonexistent, and obtaining some form of permanent documentation of the images was highly problematic. Feedback from users encouraged manufacturers to develop subsequent generations of endoscopes. Whereas the evolution of endoscopic technology did indeed require a few major improvements, such as the introduction of fiber optics and video camera capabilities, its current characteristics depend heavily on a continuous flow of refinements. These modifications have resulted in improved flexibility, maneuverability, miniaturization, and visibility and have vastly expanded the therapeutic possibilities of endoscopy.\(^5\)

This incremental, developmental activity occurs not only in industrial R&D laboratories, but in the context of clinical practice as well. For example, small departures from established practice in everyday clinical settings have yielded several important advances, a process that has produced substantial progress in surgery. But the phenomenon applies as well
to technologies that are the end result of elaborate, formal R&D processes, such as those of pharmaceuticals. When new pharmaceuticals are introduced into clinical practice, entirely new indications of use are commonly revealed. A case in point is the application of adrenergic beta-blocking drugs, one of the more significant medical innovations of our time, to new uses. These compounds were originally introduced for the treatment of two cardiovascular indications, arrhythmia and angina pectoris. Today they are used in the treatment of more than twenty diverse conditions, largely as a result of clinical discoveries made after beta-blockers were in general use. In the medical device arena, the lengthy evolution of lasers tells a similar story. Originally introduced for ophthalmologic and dermatologic purposes, the laser is now being used, or evaluated, for a wide variety of indications in gynecology, gastroenterology, oncology, thoracic surgery, and numerous other specialties.

These observations underline two critical characteristics of innovation in medicine: (1) New technologies retain a high degree of uncertainty long after their initial adoption; and (2) a close interaction between developers, often in industrial laboratories, and users is crucial to the development of new medical technologies.

One final qualification of the linear model needs to be made: The development of new interventions is influenced not only by advances in scientific and engineering knowledge (as this model implies), but also by the potential demand for particular innovations. Traditionally, physicians acting as agents for their patients have been considered the principal users by the developers of new technologies. In recent years, however, other groups—such as hospital administrators, patients, payers, and regulators such as the Food and Drug Administration (FDA)—have begun to influence the demand for technology. The preferences and rules of these various actors exert an important influence on which new technologies will be accepted into practice and how they will be used. This influence, in turn, affects the rate and direction of subsequent R&D efforts. Accordingly, this paper takes an approach to the innovation process that stresses feedback mechanisms; that is, following the development and introduction of first-generation technologies, the selection criteria and experience of these various actors may generate important new information regarding the “improvements” that need to be embodied in second- and third-generation technologies. These incremental improvements can be directed not just at enhancing performance but at redesigning to reduce costs. It is often quite possible to push these technologies in a cost-reducing direction, but financial incentives, as well as social and professional pressures in affluent societies, traditionally have pushed them in other directions. Moreover, even when these technologies are improved so that costs per unit of
“output” are reduced, the dynamics of health care are often such that these improvements in efficiency are not translated into a reduction in aggregate health care expenditures.

Whereas these observations focus primarily on the importance of (potential) demand in shaping technological change, it is also essential to examine the underlying scientific and engineering knowledge base, and the supply-side policies that stimulate its expansion. In medicine, the decades after World War II have been characterized by impressive growth in federal funding for biomedical research and medical education. The research budget of the National Institutes of Health (NIH), for example, grew from roughly $26 million in 1945 to nearly $7 billion in 1990 (both in 1988 inflation-adjusted dollars), rendering NIH the federal agency with the largest basic research budget across all sectors. This investment vastly strengthened the position of academic medical centers and stimulated the education of a growing cadre of specialists and subspecialists. In internal medicine alone, more than twelve new subspecialties formed in the past decade. This trend toward progressive specialization, as discussed below, has important consequences for the rate and direction of medical innovation.

This public commitment to biomedical research and technological advancement also is shared with the industrial producers of new technology. The U.S. pharmaceutical industry invests about 13 percent, and the U.S. medical device industry invests roughly 7 percent, of their annual sales turnover in R&D. As a result, these industries rank among the most research-intensive industrial sectors, surpassing R&D expenditures in the aircraft, electronics, and chemical industries, as a percentage of sales. In absolute terms, the industrial sector invested $12 billion in medical R&D in 1991, as compared with $10.7 billion by the federal government. Thus, both demand- and supply-side forces interact to influence the innovation process; we now discuss how the resulting innovations subsequently affect the rising costs of health care.

### Technological Change And Health Care Expenditures

Most attempts to quantify the connection between technology and rising health care expenditures employ the so-called residual approach; that is, these studies first estimate the impact of more easily identifiable factors, such as rising incomes and changing demographics. The portion of increase in health spending not accounted for is then attributed to technology. The residual approach has called attention to the role of factors that might otherwise have been ignored. However, it is of only limited use in establishing a relationship between technology and costs. Studies based upon the residual inevitably suffer from a variety of problems, the most
fundamental of which is that the reliability of any estimate derived as a residual must depend upon the confidence that all other factors, and their interaction effects, have been fully captured. Moreover, one of the main drawbacks of previous studies is that no distinction has been drawn among the intensity of use of existing technology, the introduction of new or modified technologies, and expanded applications of these new technologies. We argue that technology may contribute to the rising costs of health care through each of these three very distinct mechanisms.

**Intensity of use.** If a new technology is available for a particular indication, epidemiological studies show that it will be used at significantly different intensities in various countries, and among regions within countries. The same research also indicates that the United States exhibits a particularly intensive practice of high-technology medicine and that differences cannot be accounted for by geographical variations in disease patterns. One important explanatory variable is professional uncertainty, which is largely a result of incomplete scientific evidence concerning the effectiveness of alternative methods of intervention in everyday clinical practice. In the absence of such information, the adoption and use of technology has been shaped by a complex set of financial, professional, social, and institutional factors. Some analysts have written of a “technological imperative” that is instilled, at least in the United States, in medical students and that has been reinforced by the health care financing system. High-technology medicine is generally regarded as a source of significant professional prestige, and, in a broader cultural sense, strong social values favor its application, particularly for life-threatening disease. Furthermore, generous insurance and the presence of fee-for-service physician payment and retrospective hospital reimbursement insulated providers and patients from the immediate financial consequences of their decisions during the 1960s and 1970s. As long as new technologies were seen as offering even small health benefits compared with existing practices, and as long as third-party payers covered the incremental costs, pressures arose in the system for adoption and regular use of these technologies.

The introduction of the prospective payment system (PPS) based on diagnosis-related groups (DRGs) for Medicare recipients in the early 1980s led to great changes in hospitals’ financial incentives. Cost-plus reimbursement, which had a built-in reward for maximizing resource use, was rapidly replaced by a reimbursement method that encouraged less diagnostic testing, shorter lengths of hospital stay, and outpatient care. In addition, the recent implementation of a resource-based fee schedule for physician payment, which reduces compensation for sophisticated technology-based services, is likely to diminish the financial incentives for the application of such services. Financial incentives to reduce technology expenditures also
have been reinforced by the recent rapid growth in managed care initiatives, such as health maintenance organizations (HMOs) and preferred provider organizations (PPOs), that actively employ utilization controls. In fact, cost-conscious managed care organizations more than quadrupled their membership during the 1980s. Burton Weisbrod reports that in the 1970s HMOs had a 30 percent lower hospitalization rate than other insurance arrangements, resulting in lower costs per patient-year.\textsuperscript{14} The rate of introduction of new technologies, however, did not appear to differ at that time. Recent analyses suggest that managed care organizations may be more cautious in the adoption of particularly costly technologies, but rigorous empirical evidence is sparse.\textsuperscript{15}

The transition from a health care system in which hundreds of thousands of individual physicians are the major adopters toward one in which buying power is concentrated in a much smaller number of HMOs and other managed care organizations affects both the quantity of medical interventions and their price—and thus aggregate expenditures. Hospitals are increasingly beginning to compete on price and issue price bids for established procedures such as cardiac transplantation and coronary artery bypass grafting. Manufacturers have begun to discount prices on devices and, more recently, drugs. Unfortunately, there is very little empirical evidence that examines carefully how these changes in the delivery and financing of health care, as well as changes in the underlying technology, are affecting the price of medical interventions over time.

However, these demand-side changes interact with various supply-side factors that may affect the use of medical interventions differently. A strong relationship has been noted between the use of medical technology and the ratio of physicians to population, as well as the particular patterns of medical specialization in a country. The training, values, and interests of medical specialties and their subspecialties affect the use of medical technology; generalists, for example, have been found to use less technology than specialists for similar conditions.\textsuperscript{16} This is the more pertinent because the United States on average trains twice as many specialists as other industrialized nations (nearly 80 percent of all U.S. physicians are specialists, compared with 20 to 40 percent elsewhere). Strong institutional incentives also affect the diffusion of technological interventions. For hospitals, technology is a way to attract patients and physicians and to stay competitive. It is not surprising, therefore, that the number and distribution of hospitals is an important determinant of technology use; hospitals that have other hospitals in close proximity, which are likely to be competitors, exhibit much higher patterns of, for instance, open-heart surgery facilities.\textsuperscript{17}

\textbf{Introduction of new or modified technologies.} The above discussion still conveys a rather static view of the connections between technology
and health care expenditures; it does not do justice to the more subtle dynamics involved in technological change. The technological imperative, which implies that the medical profession is determined to exploit any new technology that comes along, is much too passive a formulation. Rather, when new technological capabilities become available, the medical profession is able to shape those technologies further either by its feedback to the R&D community or by broadening the indications of use. Moreover, other actors provide important feedback signals these days. These signals affect both the rate and the direction of innovation efforts.

Available evidence supports the view that the rate of technological change has been particularly rapid since the end of World War II. One useful index is the plethora of new drugs and devices that have been introduced: Weisbrod reports that approximately 35 percent of the 200 largest-selling prescription drugs each year are new. In 1990 nearly 5,000 new devices were introduced in the United States. The rate of innovation is sensitive to changes in the financing and delivery of health care, including the level of reimbursement (that is, price) that new interventions will be able to obtain. Percutaneous transluminal coronary angioplasty (PTCA), for example, was first assigned to a surgical DRG, which provided a much higher level of reimbursement than the procedure itself cost. This stimulated rapid adoption of PTCA and a high level of incremental innovation in PTCA catheters. By contrast, cochlear implants were placed in a DRG that covered only a fraction of the cost of the device. This led not only to underdiffusion but also to markedly reduced subsequent R&D investment. Beyond such case studies, we have not been able to find any studies that document how the recent changes in the health care marketplace are affecting the rate of innovation in particular disease categories. Somewhat more information is available about the direction of medical innovation. The direction in which developers try to move their technologies is embodied in their selection of R&D projects. As previously indicated, new technologies are often rather primitive and perform poorly at the outset, and feedback from users suggests what kinds of improvements are needed. In the past, price considerations played much less of a role in the adoption of new medical technologies than in nonmedical fields. As a result, judgments by the relevant medical specialties about a technology’s clinical performance predominated in setting the course of improvements. Feedback signals often were couched in terms of shortcomings in efficacy and safety and problems with the ease of operation—not cost reduction.

More recently, however, these signals have been changing. The growing importance of economic considerations in hospital purchasing and clinical adoption decisions is influencing technological change in the direction of developing explicitly cost-reducing technology. Less costly alternatives to
widely practiced, expensive procedures—such as coronary artery bypass surgery, transurethral prostatectomies, and cholecystectomies—have become preferred R&D targets for pharmaceutical manufacturers, which attempt to develop pharmacological alternatives, and device manufacturers, which are aiming to develop a variety of minimally invasive devices. Furthermore, interviews with pharmaceutical firms underscore that they are reallocating their R&D expenditures toward finding solutions for costly chronic care (for example, Alzheimer’s disease). By contrast, research in therapeutic categories that will be well served in the coming decade by generic drugs will be deemphasized, mainly because managed care purchasers and hospital formulary committees will encourage the substitution of generics for patent drugs. Moreover, the pressures of gatekeepers are likely to lead to decreased investment in “me-too” drugs. According to Frederick Telling, only so-called fast followers (the second or third drug in a new therapeutic class) that have a more favorable therapeutic profile or a more convenient means of administration are likely to gain relatively easy access to the market.¹⁹

Similar trends can be observed in the medical device arena. Competition on the basis of price and operating costs has begun to play a much more prominent role than in the past. Manufacturers of lithotriptors, for example, have replaced the expensive x-ray system and short-lived electrode configurations embedded in the original lithotriptor with less costly alternatives.²⁰ In the case of surgical laparoscopy, the first signs (for example, the debate on reusables versus disposables, or the emerging preferences for the much less expensive electrocautery tools over lasers) are appearing that economic considerations will increasingly influence the direction of technological change in the years to come.²¹ Interviews with device manufacturers indicate that they will reduce development of me-too technologies that do not have a clear-cut clinical or economic advantage. Research on cost-increasing—but also potentially quality-increasing—technologies (such as artificial organs) also is expected to become less attractive.

At this point, a caveat is necessary. Whereas feedback signals today may increasingly emphasize the desirability of cost reduction, we suggest that these signals are inherently ambiguous, for several (partly overlapping) reasons. First, medical research at the more purely scientific end of the spectrum contains a high degree of inherent uncertainty. Cost consequences and other features cannot be clearly foreseen. Second, even in the development and early adoption stage, many uncertainties remain. The medical applications of the laser and computer, the therapeutic applications of fiberoptic endoscopes, and the eventual uses to which adrenergic beta-blocking drugs have been put are major cases in point. Third, new medical technologies, once developed, often interact with other technolo-
gies in unexpected ways. These interactions frequently cannot be antici-
pated for the simple reason that a complementary technology may not yet
have been invented. For instance, when the earliest medical experiments
with optical fibers were undertaken in the 1950s, the laser was not yet
invented. With the subsequent perfection of laser technology and its inte-
gration into fiberoptic endoscopy, a range of new therapeutic procedures
ensued. Nevertheless, despite such uncertainties, today’s more cost-
conscious health care system is shifting the direction of medical innovation
from those interventions that are mainly driven by the search for better
clinical results to those that emphasize cost reduction.

Expanding applications. Even if current incentives are able to redirect
the innovation enterprise to the development of cost-reducing inter-
ventions, we suggest that subsequent patterns of use may nevertheless render
them cost-increasing in the aggregate. When medical specialties acquire a
new technology and the skills to use it, they are able to shape those
technologies further in ways that expand application of their services. This
expansion is encouraged by the fact that generally new technologies repres-
sent a discernible improvement in quality over their predecessors.

In the area of diagnostics, an improved technology, such as computed
tomography (CT) scanners or MRI devices, naturally suggests an expansion
of its use to broader categories of users. Modem imaging tools used to
acquire information about the brain, for example, meant that such informa-
tion could be provided noninvasively and with minimal risk. This has
dramatically broadened the technology’s indications of use, thereby prob-
ably offsetting any savings resulting from the reduction in exploratory
surgery or other interventions.

Therapeutic applications of technological capabilities that are cost re-
ducing per patient treatment but that turned out to be expenditure increas-
ing because of application to a wider population than originally anticipated
include PTCA and laparoscopic cholecystectomy. Application to a wider
population may be a consequence of two things. First, improvements in the
technology’s performance may make it applicable to patients for whom it
was not initially deemed important. PTCA, for example, was originally
estimated to be applicable to 10 to 15 percent of patients who underwent
bypass surgery (that is, only patients with a specific clinical diagnosis). As a
result of refinements in PTCA catheters, the procedure is being applied to
more complicated cases in ways that would not have been possible before,
and for patients with more advanced or unstable clinical conditions.

A further reason for application to a wider population is that before the
advent of new technologies, a sizable population may have gone untreated.
A recent article in the Journal of the American Medical Association suggests
that following the introduction of laparoscopic cholecystectomy, the over-
all level of gallbladder removals increased by no less than 60 percent in a large HMO. This is likely because the application of laparoscopic cholecystectomies has expanded from sicker to mildly symptomatic patients (suggesting that this procedure is becoming at least partly prophylactic) as well as to higher-risk patients once considered ineligible for the procedure. Thus, although laparoscopic cholecystectomies reduce unit costs by 25 percent (mostly because of shorter hospital stays), their introduction has resulted in an increase, not a decrease, in aggregate expenditures.

There is a greater elasticity of demand for medical services than is commonly believed; more precisely, a downward shift in supply may bring about an outward shift in demand, with an ultimate increase in total expenditures. When technological change not only reduces costs but also improves quality, expectations of reductions in aggregate expenditures are likely to be frustrated. Thus, cost savings obtained at a per patient level are more than offset by the increase in the use of a new technology.

Medical Specialization And Interspecialty Competition

A crucial mediating force in the interactions between technology and medical care spending is medical specialization. The behavior of practitioners is shaped by the way in which medicine is divided up into specialties and subspecialties. This is particularly true when a medical condition can be treated by different specialties, which may employ alternative or competing interventions. This means that the development and diffusion of technology do not take place in a vacuum but are influenced by the availability and development of competing technological interventions.

This issue is well illustrated by the case of gallstones, which has traditionally been defined as a surgical condition. In fact, removal of the human gallbladder, a cholecystectomy, has been for many years the most frequently performed surgical operation in hospitals. Steadily improved techniques led to an impressive decrease in mortality and morbidity rates. By the 1960s cholecystectomy was generally regarded as a safe and effective procedure, and surgeons had, in essence, a monopoly in the treatment of gallstones. Consequently, there was a relatively low level of surgical development at the time and relatively little research interest in the mechanisms underlying gallstone formation or the development of alternative treatments.

In the 1970s this began to change. Franz Ingelfinger, the father of American gastroenterology and editor of *The New England Journal of Medicine*, urged an increased research commitment. Gastroenterologists then began to develop gallstone-dissolving drugs. In a study that generated much enthusiasm at the time among clinicians, chenodeoxycholic acid (CDCA) was reported to improve the ability of bile to solubilize chole-
A New England Journal editorial indicated the emergence and intensity of interspecialty rivalries: “In the meantime our surgical colleagues can relax; their treatment of gallstones, although threatened, is not yet outmoded.” Subsequent research, however, revealed that CDCA and its successor ursodeoxycholic acid (UDCA) were of limited efficacy, and even then for only a limited patient group. Gastroenterologists in Europe and interventional radiologists in the United States took the lead in evaluating the lithotripter (originally developed for kidney stones) for gallstone treatment. At the time (mid-1980s) this intervention appeared to pose a major threat to cholecystectomy as the “gold standard” of treatment. Although we now know that biliary lithotripsy involves lengthy treatment periods, is applicable to only a few patients, and carries a high risk of recurrence, it then seemed to have potential similar to that of Tagamet and other H-2 blockers, which essentially led to the disappearance of ulcer surgery.

These competitive pressures were an important factor inducing surgeons to develop laparoscopic cholecystectomy. They adopted, and adapted, laparoscopic tools that had already been used for at least three decades by gynecological surgeons but whose application perhaps somewhat surprisingly had elicited little interest within the community of general surgeons. Surgeon-innovators and device firms presented videotapes of the first laparoscopic cholecystectomies at surgical society meetings in 1989, and afterward the procedure underwent a breathtakingly rapid rate of diffusion. More than half of the 32,750 U.S. general surgeons acquired the necessary skills within eighteen months. Whereas interspecialty competition was a powerful factor fueling the development of laparoscopic cholecystectomy, its subsequent high pace of adoption probably was stimulated more by strong patient demand. Moreover, as mentioned, following the introduction of this technique, the pool of patients undergoing gallbladder removal expanded as physicians began operating on patients they otherwise would not have considered eligible. In addition, general surgeons began to explore the value of laparoscopic surgery for a number of other indications, including hemiorrhaphies, appendectomies, and vagotomies. This is an example of new technologies shifting demand outward.

Interspecialty competition also played a role in stimulating technological change for treating angina pectoris. From the early 1900s both cardiologists and cardiac surgeons concerned themselves with its treatment. Discovery of the beneficial effect of nitrates provided cardiologists with compounds to relieve the pain that accompanies angina; in addition, during the first half of this century cardiac surgeons attempted to develop various surgical approaches for revascularizing the myocardium, especially for those with severe angina unresponsive to drugs. Mortality rates were quite high in those days, and over time these approaches were generally discarded.
In the mid-1950s growing recognition that ischemic heart disease was becoming the leading cause of death, and a major cause of illness, stimulated the interest of the pharmaceutical industry in developing new drugs. The first beta-blockers were introduced by cardiologists in the mid-1960s in Europe. U.S. regulatory policies toward beta-blockers were more stringent than in most European countries, and beta-blockers received approval for angina pectoris only in 1972. In effect, this meant that cardiac surgeons who developed new treatment procedures in the United States were able to obtain a competitive advantage over cardiologists because their surgical technology did not have to undergo regulatory approval.

Rene Favaloro, a cardiac surgeon at the Cleveland Clinic, is generally credited with the first report on coronary artery bypass grafting (CABG) in 1968. It diffused rapidly following initial discussion of the procedure at conferences and in the literature. Clinical circumstances favored swift acceptance of the operation: The condition is life-threatening, the operation made sense anatomically and physiologically, and from the outset it seemed quite effective in the relief of angina. In addition, socioeconomic conditions stimulated its use: A cadre of well-trained surgeons could offer their patients a more satisfactory procedure (especially in the absence of beta-blockers); the United States at that time had relatively few limits on expansion of its health sector; and the surgery was lucrative for both the medical profession and hospitals. By 1977 no fewer than 70,000 bypass surgeries were performed each year.

Surgery’s preeminence stimulated cardiologists to develop alternative procedures; this was reinforced by emerging limitations in the financial and surgical capacity for bypass surgery. The Swiss cardiologist Andreas Gruentzig performed the first PTCA procedure in 1977. The new technology underwent near exponential diffusion; for example, in the United States the number of procedures increased from 32,300 in 1983 to more than 200,000 in 1988. Angioplasty seemed more attractive than its costly surgical alternative, bypass surgery, which cost an average of $25,000. Whereas hospitalization for bypass surgery is around eight to twelve days, hospitalization for PTCA is generally two to three days. Not surprisingly, innovators undertook cost analyses relatively early in the development and diffusion of PTCA. G.C. Jang and colleagues estimated in 1982 that in a population of 17,000 coronary patients, substituting PTCA for CABG would result in annual savings of up to $170 million. Later studies revealed smaller cost differences, partly because the previous economic analyses had not incorporated the cost of dealing with the relatively high level of restenosis that is associated with PTCA, or the costs of emergency surgery. Nevertheless, the total cost for angioplasty at one-year follow-up was still 36 percent less than that for bypass surgery.
The rapid growth in the market for PTCA equipment stimulated a high level of developmental activity, not only in PTCA catheters but also in a wide variety of other devices, such as atherectomy catheters, stents, and lasers. These incremental improvements led to a broadening of indications for coronary angioplasty. However, at the same time that the rates of angioplasty increased, the rates of bypass surgery did not decrease, as might have been anticipated. In fact, these rates doubled during 1983-1988. In 1988 U.S. surgeons performed some 300,000 coronary artery bypass operations; spending for coronary artery bypass surgery approached $7.5 billion. This was due in part to an expansion of indications of use to the very elderly and to patients with both angina pectoris and congestive heart failure.34

These data indicate that interspecialty rivalries are a powerful factor stimulating technological change. This observation is important, because major clinical conditions, such as breast cancer and peripheral vascular disease, increasingly involve different medical specialties. For gallstones and angina, interspecialty competition led innovators to develop technologies that reduced costs per patient. Once again, however, the incentives traditionally embedded in medical care ultimately appear to have rendered them cost-increasing in the aggregate.

Policy Implications

In our view, a fruitful, more direct methodological approach to the relationship between technological change and health care expenditures would involve two ingredients. First, instead of focusing on all of medicine, one should disaggregate and focus on specific major clinical conditions. Second, within these conditions, one should examine how the three mechanisms of action identified in this paper (variations in intensity of use, introduction of new technologies, and expansion of indications of use) contribute to overall levels of utilization and price, and thus to health care costs. Such empirical analyses that “unpack” the forces underlying technological change and its relationship to health care costs are urgently needed to strengthen the basis for future policy making.

Although the lack of sufficient empirical evidence limits firm conclusions, some policy implications emerge. An important theme is how incentives in medicine have led to the use and sometimes overuse of sophisticated medical technologies, as well as to a strong commitment to their development. In the past these technologies have enhanced quality but have not reduced costs to a significant degree. If one is serious about constraining escalating health care costs, it is necessary to consider ways to both eliminate the cost-ineffective use of existing medical interventions and alter the direction of medical innovation toward the development of
cost-reducing technologies. At the same time, the difficult policy issue is to avoid reducing the incentives to innovation to such an extent that the status quo is frozen into place, mainly because much present technology is both costly and limited in effectiveness.

We believe that it is necessary to accommodate a continued commitment to innovation with the need to constrain escalating health care costs. A strong case exists for redirecting the specialist/generalist mix toward more generalists—a policy change that also figures prominently in the current health care reform debate. Of course, given the time involved in educating physicians, changing the specialty mix will not have immediate effects on technological change.

Further changes in demand-side incentives, as is being debated as global budgets, managed competition, and the like, may bring more immediate results. A notable, early example of the importance of demand-side signals can be found in the treatment of end-stage renal disease (ESRD). Following the development of dialysis, the government enacted legislation in 1972 to cover the treatment of all ESRD patients under Medicare. The considerable public expenditures that resulted from this program led the government to decrease reimbursement rates; the 1989 reimbursement rate for outpatient dialysis was 61 percent less than in 1974, when adjusted for inflation by the gross national product (GNP) price deflator. This stimulated innovation in dialysis equipment to turn in a cost-reducing direction. According to Richard Rettig and Nathan Levinsky, the equipment and supply (non-labor) component of the total cost per dialysis treatment has been reduced from roughly one-third fifteen years ago to one-fifth or less now.

It should be kept in mind, however, that today’s changing health care market sends signals to different groups that operate under different time horizons, with different priorities, and that are subject to different incentives. Indeed, the research community that is in a profit-making mode—the drug and medical device industry—is often highly sensitive to market signals. Because drug manufacturers operate under longer time horizons (their R&D pipeline is typically ten years or more) than do device manufacturers, it will take longer before the effects of redirecting R&D become visible in the pharmaceutical sector. Nevertheless, we have found preliminary evidence that both the drug and device industries are shifting their R&D priorities toward the development of cost-reducing innovations (that is, a deemphasis of me-too drugs, a shift toward ambulatory surgery, and an increase in the development of minimally invasive devices).

By contrast, NIH receives its signals not through the marketplace, but mainly through the budgetary process by which the federal government sets the level of resources for research. If the purpose is to contain escalating health care costs, it is necessary to consider seriously the reallocation of
monies within the biomedical research budget. As mentioned, NIH has the largest basic research budget of any federal agency; the substantial increase in medical research activity since World War II has provided a powerful stream of new interventions that are widely accepted as quality-improving. A cost-conscious health care system necessarily raises questions about the need to set priorities among categories of research. Given the uncertainties involved in research, is it possible to systematically identify categories that offer the best prospects for new cost-reducing technologies? A high priority for cost reduction presumably implies drastic changes in research priorities. For example, should further research in neonatal intensive care units continue to receive a higher financial priority than research that explores ways to persuade more women to use prenatal care services? Despite the difficult and contentious nature of such an undertaking, it is necessary to begin exploring the promise of different categories of research, with the goal of reducing medical expenditures and improving clinical results.

A second observation about the allocation of the R&D budget concerns assessing costs and benefits. The United States over the past forty years has been extraordinarily successful in producing new knowledge. At the same time, there has been a deplorable failure in assessing the benefits and costs of putting this knowledge to use. A 1984 study by the Institute of Medicine (IOM) found that U.S. expenditures for assessment activities were a mere 0.3 percent of total health care expenditures. In comparison with the $10 billion research budget of NIH (mostly spent on basic research), the budget of the Agency for Health Care Policy and Research (AHCPR) - the agency with a mandate for health services and outcomes research - was only around $120 million in 1992. Thus, there is a clear need for more appropriate levels of funding for outcomes research. With improved levels of funding, new technologies need to be assessed as early as possible. A caveat, however, is essential: Improvements in medicine are mostly incremental and part of a continuous process of numerous small-scale advances. Consequently, the manner of use, the clinical results achieved, and the resource costs associated with technological interventions change continually. This argues for an evaluative approach that emphasizes periodic reassessment of an intervention during its life cycle.

Of course, shifting the direction of innovative activity and increasing efforts to assess the benefits and costs of new and existing technologies is not going to matter unless the results of these assessments are integrated into clinical decision making. The way in which a new technology ultimately will affect costs depends on the manner in which it is incorporated into the larger system of medical care - how the profession chooses to use it and to modify it. In addition to their role in developing new medical interventions (often in collaboration with industrial firms), physician-
innovators tend to find new patient indications for existing interventions. Although some of these new clinical uses may well be cost-effective, we suspect that the broadening of indications ultimately leads to marginally beneficial applications that raise overall spending levels. An important policy challenge in the years ahead is to tie the results of assessments more strongly to clinical decisions, either directly (through feedback to the clinical community) or indirectly by integrating them with regulatory or reimbursement mechanisms. Ultimately, even the most sophisticated techniques of technology assessment are of little operational significance unless they provide a basis for restricting expenditures on medical interventions that are of minor social benefit.

We acknowledge the numerous suggestions received from Henry Aaron, Richard Rettig, and Scott Stem in the preparation of this paper, even though we have not managed to absorb all of their excellent advice.

NOTES

21. Gelijns and Rosenberg, “From the Scalpel to the Scope.”
26. Ingelfinger, “Digestive Disease as a National Problem.”
36. Ibid.