I

In Health Care, Geography Is Destiny

Early in my career, I was hired as director of a federally sponsored program whose goal was to ensure that all Vermonters had access to recent advances in the treatment of heart disease, cancers, and stroke. As part of the program, my colleagues and I developed a data system that we thought would help us identify which Vermont communities were underserved, and thus in need of the program's help. As the results came in, however, rather than evidence for underuse (i.e., patients not getting care they needed), we found extensive and seemingly inexplicable variation in the way health care was delivered from one Vermont community to another. In Stowe, for example, the rate of tonsillectomy was such that by age 15, about 60% of children were without tonsils, while in the bordering town of Waterbury, only 20% had undergone the surgery by that age. Among communities, the chances that a woman would have her uterus surgically removed varied by more than fourfold, and the rate of gallbladder surgery varied by more than threefold. Rates of hospitalizations for a host of different medical conditions also varied in ways that made little sense; on a per capita basis, patients were hospitalized in Randolph two times more often for digestive disease than in Middlebury and three times more often for respiratory disease.

These are just a few examples of the chaotic patterns of utilization and practice our data uncovered—variations that challenged the very premise of the program I had been hired to direct. The rates of hospitalization and surgery appeared to be unrelated to illness or other patient-based factors, and thus the variation was at odds with the conventional wisdom that medicine was driven by science and by an understanding of patient desires and preferences. The data also challenged the assumption that the supply of medical resources and the capacity of the health care system were regulated either by a central professional consensus on the need for medical care, and its effectiveness, or by the invisible hand of the market. It became clear that the amount of care Vermonters received depended on where they lived and on the physicians and hospitals they used.

Over the years since that time, my colleagues and I have pursued the study of practice variation in many places, using a variety of methods, and the Vermont findings have been widely confirmed. Unwarranted variation in health care delivery—variation that cannot be explained on the basis of illness, medical evidence, or patient preference—is ubiquitous. Moreover, as I argue in this book, an understanding of the causes of unwarranted variation has important and sometimes surprising implications for today's debate over health care reform. Most analysts of health care reform expect huge increases in spending once the uninsured gain coverage and begin to consume more health care services. But the understanding I have gained from the study of the practice variation phenomenon provides a counterintuitive, maybe even shocking, prediction: given the important role that the supply of resources plays in determining utilization of medical care, increasing the insured population will have a much smaller impact on the trend in overall health care costs than estimated, provided that the capacity of the health care system is not increased.

Another prediction that emerges from an understanding of practice variation is that controlling costs will not necessarily require rationing—if by "rationing" we mean the withholding of care that patients want, and that is effective in improving outcomes. The studies reviewed in this book show that much of health care is of questionable value and that informed patients often prefer a form of treatment other than the one their physicians actually prescribe. Indeed, when offered a clear explanation of the treatment options, informed patients often choose the less invasive treatment, resulting in a decline in the use of elective surgery and certain cancer screening tests. Moreover, more care is not necessarily better, at least when it comes to managing chronic illness. Care coordination and intelligent management of patients over the course of their illness, which typically lasts until death, count far more than simply providing more medical services. Some of our most respected health care providers—for example, the Mayo Clinic, the Geisinger Clinic, and the Cleveland Clinic—provide high-quality care at a

much lower per capita cost than most other providers. If the rest of the nation were equally efficient, we could shave 30% to 40% off the cost of caring for Medicare's chronically ill patients.

If, as I recommend in this book, health care reform concentrates on four goals, the quality and value of care will increase and growth in health care costs will likely decrease. Those goals are as follows:

1. Promoting organized systems of health care delivery
2. Establishing informed patient choice as the ethical and legal standard for decisions surrounding elective surgeries, drugs, tests, and procedures, and care at the end of life
3. Improving the science of health care delivery
4. Constraining undisciplined growth in health care capacity and spending

These are strong conclusions, ones that policy makers should not ignore given today's economic realities. They are supported by a growing body of evidence drawn from practice variation studies and from interventions to improve the scientific basis of clinical decision making and promote informed patient choice. An important goal of this book is to make this complicated and interconnected body of research accessible to a broad audience, including policy makers, health care providers, students, patient advocates, and, I hope, patients and families.

Epidemiology of Medical Care

My understanding of practice variation is based primarily on evidence from "medical care epidemiology," studies that use routinely collected data (primarily from insurance claims) to conduct what we have dubbed "small area analysis of health care delivery." An important feature of the small area methodology is that it is *population-based:* it studies the use of health care services among populations living within the geographic boundaries of "natural" health care markets. Our Vermont studies were extended to Maine and eventually replicated throughout New England and in Iowa. In the early 1990s, anticipating that the Clinton health plan and its provision for regulating health care at the regional level would become law, the Robert Wood Johnson Foundation provided us with the funds to use claims data from the Medicare program to develop a body of data that would provide feedback to both Medicare administrators and providers, and a means of bringing practice variations to the attention of those who would implement reform. By the time it became clear that the Clinton plan had failed, we had completed much of the research but had lost our primary customers.

The failure of the Clinton plan led to the establishment of the Dartmouth Atlas Project. Rather than use the remaining funds solely for research, Dr. Steven Schroeder, then president of the Robert Wood Johnson Foundation, and James Knickman, its vice president, encouraged us to stick to the plan to provide feedback but to target a wider audience, in the hope that information on local and regional practice variation would focus attention on the need to reduce it. With support from several foundations, we have continued to analyze the care delivered to Medicare enrollees and have made the results available on the Dartmouth Atlas website (www.dartmouthatlas.org). Most of our published reports (and much of the data I use in this book) compare the geographic practice patterns among the Medicare enrollees living in 1 of 306 hospital referral regions (Box 1.1).

Box 1.1. *The Geography of Health Care in the United States*

The use of health care resources in the United States is highly localized. Most Americans use the services of physicians whose practices are nearby. Physicians, in turn, are usually affiliated with hospitals that are near their practices. As a result, when patients are admitted to hospitals, the admission generally takes place within a relatively short distance of where the patient lives. This is true across the United States. Although the distances from homes to hospitals vary with geography—people who live in rural areas travel farther than those who live in cities—in general, most patients are admitted to a hospital close to where they live to obtain an appropriate level of care.

The Medicare program maintains exhaustive records of hospitalizations, which makes it possible to trace the patterns of use of hospital care. (Research shows that the pattern of use by patients in the Medicare program is more or less similar to that of younger patients.) In the Dartmouth Atlas Project, 3,436 geographically distinct hospital service areas in the United States were defined. In each hospital service area, most of the care received by Medicare patients is provided in hospitals within the area. Based on the patterns of care for major cardiovascular surgery and neurosurgery (which are generally provided at tertiary care hospitals), hospital service areas were aggregated into 306 hospital referral regions. (Details on how hospital service areas and referral regions are defined are given in the Appendix, and maps showing their location are available in the Methods section of the 1999 Dartmouth Atlas.[1])

It is important for the reader to keep in mind that the comparisons are *population-based.* We look at what happens to groups of patients, not individuals, and we compare what happens to those groups living in different parts of the United States. In calculating the numerator for a population-based rate, all medical services are counted, *regardless of where in the United States care was obtained.* For example, if a resident of the Fort Meyers region goes to a hospital located in the Miami region to get surgery, the procedure is counted as a service delivered to the population living in Fort Meyers. Looking at populations in this way allows us to document large differences in the way care is delivered by different health care providers, but it also offers individuals a way to understand what might happen to them, depending on where they live and where they go to get their care. In making population-based comparisons—whether among regions or populations loyal to a given hospital—the rates are adjusted for differences in important characteristics of the population that influence the use of health care, such as age, gender, race/ethnicity, and, when possible or appropriate, type and severity of illness. For details on the methods used in this book, please consult the Appendix.

One of the more powerful and inescapable conclusions that has emerged from our research is that physician behavior is behind much of the variation. I do not mean that all, or even most, physicians are cynically rubbing their hands together every time a patient walks in the door, thinking of ways to deliver more care, and thus make more money. On the contrary, most physicians are simply trying to do the best job they can to care for patients. Nonetheless, physicians practice in a particular context—in a local market with its own complement of health resources, including the supply of hospital beds and physicians. It is physicians who exert the greatest influence over demand—or really, utilization—because patients traditionally delegate

Figure 1.1. The 306 Dartmouth Atlas Hospital Referral Regions.

decision making to them under the assumption that doctors know what is best. Physicians thus control the majority of decisions made in medicine, most of which do not necessarily put money in the physician's pocket. The most costly decisions are those governing the use of acute care hospitals.

Categories of Care

My understanding of the role that physicians play in influencing demand has been greatly facilitated by the realization that the causes and the remedies for unwarranted variation differ according to three categories of care: effective or necessary care, preference-sensitive care, and supply-sensitive care. Until very recently, policy makers have concentrated almost exclusively on what I call the "*effective care*" or "*necessary care*" category—services that, on the basis of reasonably sound medical evidence, are known to work better than any alternative, and for which the benefits of treatment far exceed the side effects or unintended consequences. In other words, effective care includes any treatment that all eligible patients should receive. Demand for effective care is defined and limited by medical science—by objective information, "high-quality" information about the outcomes of treatment and evidence-based clinical guidelines that identify which patients stand to benefit.

For effective care, the problem is underuse—the failure to provide care for patients who should, but did not, get the required treatment. Examples of underuse include failure to provide immunizations to young children or lifesaving drugs to patients with heart attacks. Efforts by policy makers to increase the use of such effective care include monitoring of performance, publication of quality reports on the Internet (such as Medicare's Hospital Compare! Website), and Medicare's Pay for Performance program, which rewards providers who achieve high rates of use of effective care and penalizes those who achieve low quality scores.

Although it is important to reduce underuse, it does not account for much of the overall variation in Medicare spending. Even when one includes the inpatient costs for conditions and treatments for which there is no alternative to hospitalization (e.g., hip fractures and surgery for colon cancer), spending for effective care seems to account for no more than about 15% of total Medicare spending. Ironically, our research shows that greater supply of physicians and greater total Medicare per capita spending are not associated with less underuse of effective care.

A second category of care that varies is elective, or "preference-sensitive" care, interventions for which there is more than one option and where the outcomes will differ according to the option used. This category, which accounts for about 25% of Medicare spending, includes elective surgery, for example, and such cancer screening tests as mammography and the prostate specific antigen test. The treatment of early-stage breast cancer provides a good example of preference-sensitive surgery. For most patients, the options include lumpectomy, or local excision of the cancer, and mastectomy, the complete removal of the breast. The two are equivalent in terms of impact on reducing mortality but have very different impacts on the quality of life; thus, the decision as to which treatment is right for the individual patient should depend on the patient's preference. But for reasons described in this book, because patients delegate decision making to doctors, *physician opinion* rather than *patient preference* often determines which treatment patients receive. I argue that this can result in a serious but commonly overlooked medical error: operating on the wrong patients—on those who, were they fully informed, would not have wanted the operation they received. Figure 1.2 is a graphic representation of the various forces I will discuss in this book that come into play for preference-sensitive care when patients delegate decision making to their physicians.

Finding a remedy for unwarranted variation in preference-sensitive care requires a concentrated, ongoing effort to reduce scientific uncertainty about the outcomes of various treatments. But evidence-based medicine is only

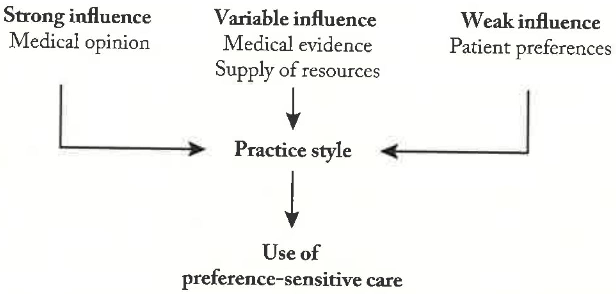

Figure 1.2. A model of preference-sensitive care under delegated decision making.

part of the answer. The more fundamental reform must involve a shift in the culture of medicine—a change in the doctor-patient relationship that reduces the influence of medical opinion and enhances the role of patient preferences in determining the utilization of preference-sensitive care. This democratization of the doctor-patient relationship requires replacing delegated decision making, and the doctrine of informed consent, with shared decision making and informed patient choice. My book will argue that establishing evidence-based medicine and informed patient choice are feasible as well as necessary goals for health reform.

The third category of care is what we have come to call "supply-sensitive care." It differs in fundamental ways from both effective care and preference-sensitive care. Supply-sensitive care is not about a specific treatment per se; rather, it is about the frequency with which everyday medical care is used in treating patients with acute and chronic illnesses. Here I am talking about physician visits; referrals for a consultation, home health care, and imaging exams; and admissions to hospitals, intensive care units (ICUs), and skilled nursing homes. The physicians whose decisions determine the frequency of such care are not usually surgeons—they are mostly primary care physicians and medical specialists.

This category, which accounts for roughly 60% of Medicare spending, may be difficult to grasp because it runs counter to the widespread belief that medical interventions are driven by explicit medical theories and scientific evidence. Most of us, including most doctors, believe that a physician makes decisions such as when to schedule a patient with diabetes for a follow-up visit, for example, or when to hospitalize a patient with chronic heart failure, or when to call in an infectious disease specialist for a patient with a fever, on the basis of medical science, augmented by some combination of

experience and wisdom. As it turns out, medical science is virtually silent on such matters.

There is another factor that influences such decisions. As Figure 1.3 illustrates and the book will demonstrate, physician decisions regarding supply-sensitive care are strongly influenced by the capacity of the local medical market—the per capita numbers of primary care physicians, medical specialists, and hospital or ICU beds, for example. (In the jargon of economics, the market is in disequilibrium—supply pushes demand or utilization.) This may seem deeply counterintuitive, and the effect of supply on professional behavior by and large goes unrecognized by physicians, who are unaware of the effect that capacity has on their decisions. But in the absence of a constraining professional consensus on best practices, and under the cultural assumption that more care is better care, available resources are used up to the point of their exhaustion. Moreover, patients who live in regions of the country where per capita supply of resources is high have no way of knowing that they are destined to spend more days in the ICU, for example, days that they probably would not have spent had they lived in a region of the country where the per capita supply of ICU beds was less.

Remedying variation in supply-sensitive care requires coming to terms with the "more care is better" assumption. Are physician services and hospitals in high-cost, high-use regions overused? Or is valuable care being rationed in regions with low rates of use, even though physicians and their patients are unaware of it? Beginning with the early studies in Vermont, extended to comparisons between Boston and New Haven, and now accomplished on a national scale as part of the Dartmouth Atlas Project, our studies consistently show that more care is not necessarily better.

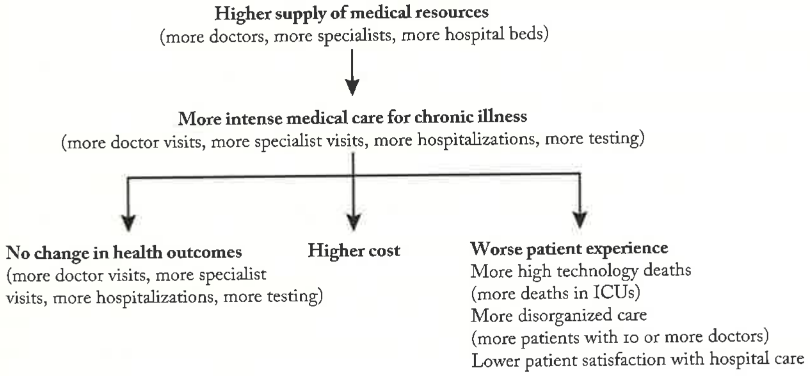

Figure 1.3. A model of supply-sensitive medical care.

Finding a remedy for unwarranted variations in supply-sensitive care requires improving the science of health care delivery—converting the "black box" of supply-sensitive care into evidence-based care that is effective or preference-sensitive, thus reducing the power that capacity exerts on use. As this book will argue, this is particularly important for patients with severe chronic illness, whose care today is primarily driven by local capacity of the delivery system, not by the wishes of patients and their families, particularly during the last two years of life. Reducing the overuse of supply-sensitive care will also require organized systems of delivery, capable of managing the care of a population of chronically ill patients over time and across locations of care, and adjusting capacity to reflect medical evidence and patient preferences. The good news for the health care economy is that, compared with most providers, organized systems of delivery are relatively efficient. They use fewer resources (and spend less) in serving their chronically ill patients and, by available measures, achieve high-quality care and satisfied patients. The bad news is that the United States does not have enough of them. My book will argue that conducting the necessary research and promoting the growth of organized systems are necessary goals for health reform.

I have organized this book to highlight the importance of, first, preference-sensitive care and, then, supply-sensitive care, which together make up about 85% of Medicare spending. The following chapter tells the story of my first encounter with practice variation in Vermont and how the extent and magnitude of the variations we uncovered challenged me to reconsider some basic assumptions about how health care markets worked. Chapters 3 through 7 are devoted to unwarranted surgical variation, to understanding the patterns of variation, the role of medical opinion as a cause of variation, the role that supply of resources sometimes plays in decisions about preference-sensitive treatments, the importance of reforming the doctor-patient relationship to ensure that patient preferences play a part in determining when surgery is necessary, and to a report on the research project we undertook to learn how well surgical treatments work and to help patients make decisions on the treatment they want. I argue that this project, undertaken over more than ten years, provides a cogent model for how the science of health care delivery can reduce uncertainty, clarify the importance of patient preferences and address significant flaws in the market for health care services.

Chapters 8 through 12 are dedicated to understanding supply-sensitive care. I review the effect that supply exerts on care intensity—the frequency of hospitalization, for example—for those with chronic illness and the evidence that greater care intensity is not driven by differences in illness and that greater intensity is not producing better outcomes, and I make estimates of the waste from the overuse of supply-sensitive services. Even academic medical centers, "America's Best Hospitals," are shown to vary widely in their treatment patterns, much like other hospitals. However, organized systems of care—multispecialty group practices and integrated hospital systems—are generally more efficient: compared with most providers in the United States, they use fewer resources to deliver equal, often higher-quality care. When we use the per capita resources of organized care systems as benchmarks for the rest of the country, I see a glimmer of hope with regard to controlling health care spending. The efficiency achieved by these organized practices suggests that the nation already has more than enough resources and spends more than enough to care for all Americans, provided we can "reengineer," or transform, the rest of the system so that it looks more like those of organized group practices and less like the disorganized, fragmented, inefficient delivery nonsystem that currently exists.

Chapters 13 through 15 focus on the four goals of health care reform that I have set. I suggest strategies for improving the science of health care delivery, promoting the growth of organized systems of care, and establishing informed patient choice as a standard of care. But I want to be clear. While we urgently need to reform the health care delivery system, the nation cannot depend on the reengineering of clinical practice as the primary strategy for achieving the fourth goal: constraining undisciplined growth in capacity and out-of-control expansion in health care spending. Reducing unwarranted variations requires a painful transition from today's chaotic, disorganized care to systems of organized care and a cultural change from patient dependency on the authority of the physician to the democratization of the doctor-patient relationship. How long this will take simply cannot be predicted, but it will likely take years before these reforms can be expected to play a significant role in controlling the growth of costs. In the meantime, unless specific steps are taken to counter the dynamics of growth, the health care spending bubble will continue to expand, further threatening the national economy and limiting our options for designing our future. It is up to policy makers to take the necessary steps. In Chapter 15, I outline five steps that can be taken to place limits on capacity and spending, and buy time for reform to take hold.

* * *

The final chapter is a summing up of the challenges we face. An Epilogue looks at the prospects that federal legislation will advance my goals for health care reform.