

INTRODUCTION TO Public Health

Third Edition



Mary-Jane Schneider

Health Services Research: Finding What Works



Will Back Surgery Help?

In the late 1960s and early 1970s, the medical establishment was shaken by a number of reports that documented wide variations in the way physicians treated their patients for common health problems. One study found that in Morrisville, Vermont, nearly 70 percent of the children had their tonsils removed by the time they were 15 years old, whereas in nearby Middlebury, only 8 percent of children underwent the operation. Another study in Iowa reported that more than 60 percent of the male population of one community had their prostate glands removed by age 85, whereas the rate was only 15 percent in another area. And the rates at which women underwent hysterectomy varied from 20 percent in one part of Maine to 70 percent in a city less than twenty miles away.¹

The reasons for these differences were unclear. The populations of the comparison communities were not substantially different from one another. There was no reason to believe that the residents of one community were sicker than those of another or that their insurance coverage was more comprehensive. It seemed obvious that these procedures were being overused in some geographical areas or underused in others. However, the studies could not determine which was true or decide what the appropriate use rates should be.

This method of examining medical practice, known as small-area analysis, has been applied over the past three decades to a broad range of medical practices and procedures. Repeatedly, wide variations have been found, with no apparent reason for the differences in practice. In 1996, Dr. John Wennberg, a professor at Dartmouth Medical School and a pioneer in the field who had conducted the studies in Vermont, Maine, and Iowa, published the *Dartmouth Atlas of Health Care*, an analysis of 1993 Medicare data.² (Since the Medicare program maintains files on everything it pays for, including services to virtually all Americans 65 and older, it provides valuable data for this kind of research.) All over the country, variations occur in treatments for prostate cancer, breast cancer, heart disease, and many other common conditions. Rates of surgery for low back pain were seven times higher in Provo, Utah, than in Kingsport, Tennessee. Breast-sparing surgery for women with breast cancer ranged from just over 1 percent in Rapid City, South Dakota, to 49 percent in Elyria, Ohio.

Small-area analysis called attention to the lack of scientific evidence on which doctors and patients base decisions about how various medical conditions should be treated. The surprising results of the early studies were part of a new field of research—health services research. This research attempts to understand the reasons for the observed variations in medical practice and to determine, from observations of the everyday practice of medicine, what treatments lead to the most desirable outcomes. Health services research studies the effectiveness, efficiency, and equity of the healthcare system. It is a way of trying to assess the quality of medical care. This research also may lead to insights on how to control costs and improve access.

Reasons for Practice Variations

A number of explanations have been suggested for variations in medical practice, most of which can be tested, and most of which can be shown to play a role in the observed differences. It is clear that the variability in the use of different treatments reflects the degree of uncertainty facing physicians regarding their relative efficacy. Variations in practice are far greater for some medical conditions than for others. For example, most physicians agree that surgery is the appropriate treatment for appendicitis and broken hips. Correspondingly, the geographic variability in the treatment of those conditions is much smaller than the variability in rates for tonsillitis and disorders of the uterus, on which there is much less evidence about when surgery is needed.

In many cases, doctors are unaware that their way of treating a condition is unusual, and they will change their patterns of practice when presented with evidence that they are deviating from the norm. In the early 1970s, Wennberg confronted the physicians of Morrisville, Vermont, with data showing that they were doing tonsillectomies far more frequently than other doctors in the state. The Morrisville physicians reconsidered the indications for the procedure, instituted a policy of obtaining second opinions before deciding on surgery, and ended by reducing the tonsillectomy rate to less than 10 percent of what it had been.¹

It is easy to suspect that inappropriate use of invasive procedures is responsible for the observed variations in the frequency with which they are done. While this suspicion is supported by the Vermont experience in reducing tonsillectomy rates, other studies have found that inappropriate use explains only a small part of the wide variability observed for many procedures.

In one small-area study of three procedures commonly done on Medicare patients, panels of expert physicians examined the files of a random sample of patients who had undergone each procedure. The experts compared the indications for the procedure in a high-use area with those in a low-use area. They were asked to determine, for each patient, whether the decision to do the procedure was appropriate, equivocal, or inappropriate. Coronary angiography—used to identify blockages in the blood vessels of the heart—was performed more than twice as frequently in the high-use area as in the low-use area. Yet even in the high-use area the experts considered it inappropriate in only one sixth of the cases. Carotid endarterectomy (CEA), which had an almost four-fold variation in frequency, was the procedure most often judged inappropriate. A risky procedure intended to remove blockages in the arteries that carry blood to the brain, CEA was deemed inappropriate in about one third of the cases done in the high-use area. However, the procedure was considered by the experts to be inappropriate almost as frequently in the low-use area.³

This evidence suggests that, for many medical conditions, more than one response may be appropriate. When faced with a patient suffering from a specific illness, one physician may prefer conservative treatment using drugs and “watchful waiting,” while another physician may believe that immediate surgery is indicated. These opinions tend to be shared by the physicians within a community. Wennberg has called these differences the “practice style” factor. For most of the conditions in question, there was not enough scientific evidence to determine which treatment yields a better outcome for the patient. In many cases, the choice of treatment involves weighing benefits against risks, a trade-off that different patients might evaluate differently if they are given the opportunity to choose.

The high variability and frequent inappropriate use of CEA, together with the high risks from the procedure, inspired several large randomized controlled trials, involving over 10,000 patients, to clarify the indications for and efficacy of CEA. The trials demonstrated that, among carefully selected patients and surgeons, the procedure reduced the risk of stroke and death compared with medical therapy alone. In a later analysis to determine whether the evidence provided by the trials changed medical practice, researchers in New York State conducted a cohort study of all Medicare patients who had had a CEA over an eighteen-month period in 1998 and 1999. The results were a great improvement over the earlier study: overall 87.1 percent of the procedures had been done for appropriate reasons; 4.3 percent had been done for uncertain reasons; and 8.6 percent had been done for inappropriate reasons.⁴

As for coronary angiography, another procedure studied earlier, no such randomized trials have been done to determine appropriateness. It is still a high-variability procedure: a recent study comparing rates in different states found a 53 percent higher rate in Florida than in Colorado. The rate depended in part on the density of specialists in the area.⁵

The Field of Dreams Effect

One factor that has consistently been shown to influence practice styles is the availability of services in a community, as shown in the rates of coronary angiography discussed above. The presence of a greater number of surgeons is accompanied by the performance of a larger number of surgeries; higher numbers of hospital beds lead to higher rates of hospitalization. This effect was dramatically illustrated in Maine during the early 1980s, when two neurosurgeons moved to a community and devoted themselves to performing laminectomies—disc surgery for low back pain. The number of laminectomies for the whole state nearly doubled as a result of the work of these two surgeons, although only 20 percent of the population of Maine lived in that community and the adjacent referral area.⁶ This high rate of surgery, like the tonsillectomies in Vermont, came down after the surgeons were confronted with data on practice patterns in other communities.

Research has consistently demonstrated an influence of supply on usage when hospital beds are concerned. A study done in the 1980s comparing Boston, Massachusetts, with New Haven, Connecticut, found that Boston had 4.5 hospital beds per thousand people, whereas New Haven had only 2.9 beds per thousand. Approximately the same percentage of beds was filled in the two cities, meaning that the population of Boston was hospitalized at a higher rate than that of New Haven. When Wennberg and his colleagues interviewed physicians in the two cities, they found that New Haven doctors were not purposely trying to ration care and that neither group of doctors was aware that they hospitalized patients more or less frequently than average.⁶ Mortality rates and other measures of quality of care were almost the same in the two cities. In the 1990s, when managed care came to both cities, hospitalization rates fell, but by about the same percentages, so that Boston doctors still hospitalized their patients much more frequently than New Haven doctors.⁷

The Dartmouth researchers' analysis of Medicare data found that the number of hospital beds in a community significantly influences the kind of care received by dying elderly people.⁸ Medicare patients in New York City, Newark, New Jersey, and Memphis, Tennessee, are much more likely to spend their final days in a hospital, often in an intensive-care unit, than elderly patients in Portland, Oregon, or Salt Lake City, Utah, who are more likely to die at home. Based on 1994 and 1995 data, the rates at which Medicare patients die in the hospital correlate closely with the number of hospital beds per thousand residents in their community. Researchers call this correlation the "Field of Dreams Effect," after the line in the 1989 movie about a baseball field: "If you build it, they will come."

While there is little evidence to show that patients are helped or harmed by the more intensive care they receive in Boston, Miami, and other high-use areas of the country, the differences in use have a major impact on medical care costs. For example, the average hospital bill for each Medicare enrollee's final six months of life is \$16,571 in the New York City borough of Manhattan, as opposed to an average of only \$6793 in Portland, Oregon.⁸ In the Boston–New Haven comparative study, Boston's per capita hospital expenditures were about double those of New Haven.⁹ Wennberg calculated that overall hospital expenditures for the 685,000 residents of Boston were \$300 million higher in 1982 than they would have been if the usage rates of New Haven applied.

Wennberg does not specifically argue that conflict of interest or pecuniary motives enter into decisions that determine use rates of medical services. However, many studies suggest that financial considerations may enter into some physicians' medical decision making. For example, there is evidence that when physicians stand to profit from the performance of diagnostic tests, they are much more likely to order such tests. Until the practice was outlawed by Congress, physicians who owned an interest in clinical laboratories were more likely to refer patients for laboratory tests than similar physicians who referred patients to labs in which they

had no financial interest.¹⁰ Similarly, physicians who own diagnostic imaging equipment are more likely to use it than comparable physicians who must refer patients elsewhere for such examinations.¹¹ A recent surge in complex spinal-fusion operations has been linked to the high rates Medicare will pay to surgeons and hospitals, although there is no evidence that the procedure is more effective at curing back pain than laminectomies or even less invasive approaches.¹²

Outcomes Research

As we have seen, variations in medical care are greatest for medical conditions for which the least is known about the effectiveness and appropriateness of various diagnostic and treatment approaches. The solution to the uncertainties raised by small-area analysis, therefore, is to study outcomes of these various diagnostic and treatment approaches in order to determine what works. Many policy makers believe that such research will allow the development of guidelines for medical practice, leading not only to more effective medical care but also to cost savings through the elimination of unnecessary care.

Outcomes research is the epidemiologic study of medical care. Whereas epidemiology usually examines the disease-causing effects of exposure to agents such as viruses and toxic chemicals, medical care epidemiology examines the health effects of exposure to medical interventions. Controlled clinical trials are one form of medical care epidemiology, but there are practical, financial, and ethical barriers that prevent conducting controlled trials aimed at answering many important questions about medical care. Outcomes research collects and analyzes data generated by the everyday practice of medicine in order to reach conclusions on benefits and risks of various interventions for various types of patients.

One of the early questions John Wennberg's group looked into was prostatectomy, the surgical removal of men's prostate glands. It was a high-variation procedure; in some parts of Maine, 60 percent of the men had their prostates removed by age 80; in other parts, less than 20 percent had.¹³ The procedure is used as a treatment for cancer of the prostate and for benign prostatic hyperplasia (BPH), a common condition in older men that causes difficulties with urination. Other treatments are available for both conditions, including watchful waiting, since many cases of prostate cancer never progress to become life threatening. For BPH, proponents of the surgical procedure argued that it could reduce symptoms and improve the quality of men's lives. Skeptics point out that surgery often has unwelcome side effects.

Wennberg and his colleagues conducted a major analysis of Medicare records to determine outcomes of surgery for BPH. They found that published reports significantly overstated the benefits of prostatectomy and understated the complications. Although only about 1 percent of men died in the hospital, 2 to 5 percent of the patients died in the weeks following the surgery. Moreover, within four years of the surgery, almost half of the patients had required further

treatment for urinary tract problems. After eight years, about one in five had needed a second prostatectomy.⁶ Having the surgery did not increase life expectancy, and the effect on quality of life was mixed: it improved urinary tract symptoms, but it had a negative impact on sexual function.¹³

The results of these studies indicate a need for better informing patients about their choices and about the probable outcomes of each choice.¹⁴ Feelings about symptoms, willingness to accept risks of the surgery, and personal assessment of the possible outcomes vary substantially among individuals. Research base on outcomes should enable these patients to make informed decisions based on their own values. As Wennberg has pointed out, "current rates of use of invasive high-technology medicine could well be higher than patients want.... Given an option, patients will on average select less invasive strategies than physicians."^{7(p.1203)} Effective drug therapies have been developed for BPH, and the number of surgeries performed for this condition declined in the 1990s, perhaps due in part to evidence contributed by outcomes research.¹⁵

The number of prostatectomies for cancer has increased, however, due in part to the development of a new screening method that became widely used in the 1990s. The test measures prostate-specific antigen (PSA) in the blood, levels of which have been correlated with the presence of cancer. However, low-grade prostate cancer is very common in older men, and many cases never progress to cause a problem. The follow-up testing and treatment of men whose PSA levels are elevated is invasive and may have undesirable side effects. The problem with the use of PSA screening is that there is no evidence that it reduces mortality from prostate cancer.

In a study conducted by the Dartmouth researchers, Medicare data were used to compare two cohorts of men who lived in areas with different practice patterns for screening and treatment. In the Seattle–Puget Sound area, men were tested at a rate 5.39 times the rate in Connecticut. The researchers found that more than twice as many men in the Seattle area, compared with Connecticut men, were subjected to biopsies of the prostate to confirm the presence of cancer. The Seattle area men were over five times more likely to have a prostatectomy than the Connecticut men. However, after eleven years of follow-up, there was no significant difference in the mortality rates from prostate cancer between the two groups of men.¹⁶ This finding was confirmed in 2009 with the publication of results from two clinical trials that followed a total of 259,000 men in the United States and Europe for seven to ten years. In both trials, men were randomly assigned to groups with and without PSA screening, and there was little difference in mortality between the two groups. Whether a benefit will show up after a longer period of follow-up remains to be seen.¹⁷ The problem with finding prostate cancers through screening is that there is no good way to determine which ones are likely to progress rapidly and cause harm and which are indolent and can be left alone.

Inspired in part by Wennberg's work, Congress in 1989 established the federal Agency for Health Care Policy and Research (AHCPR), hoping that studies such as those on BPH would encourage a reduction in high-technology medicine and save money on medical costs, especially for Medicare and Medicaid. The agency was mandated to examine the reasons for the wide variations in healthcare practices around the country, develop guidelines for treatment, and find effective ways to disseminate its research findings and guidelines.¹⁸ However, the Agency—and Congress—discovered to their surprise that the research results were not always welcome.

One of the health conditions that the AHCPR tackled early was low back pain. It is a widespread problem, ranking second only to the common cold as a reason that people go to the doctor. Treatment of low back problems costs over \$20 billion a year in the United States. Surgery for low back pain is a high-variability procedure, ranging from a low in the Northeast to a rate in the Northwest that is more than three times higher. The guidelines developed by AHCPR's panel of experts and released in December 1994 recommended treating most acute, painful low back problems with nonprescription painkillers and mild exercise, followed in about two weeks by conditioning exercises. Surgery benefits only about 1 in 200 people with acute low back problems, according to the chairman of the panel, a professor of orthopedic surgery at the University of Washington School of Medicine.¹⁹

Back surgeons responded with rage and political action. With the Republican Congress intent on budget cutting in 1995, legislators were sympathetic to claims by the back surgeons' lobbying group that AHCPR was a waste of money, that the government should not be telling doctors how to practice medicine, and that the agency should be eliminated.²⁰ Defenders of the AHCPR pointed out that the guidelines could save billions of dollars and accused back surgeons of merely trying to protect their incomes. When the federal budget was finally approved that year, AHCPR had survived, although its budget was cut substantially. Its leaders decided that developing clinical guidelines was too dangerous politically, but the agency continued collecting evidence that allowed other organizations to do so, and it maintains a national clearinghouse of evidence-based clinical guidelines developed by other organizations. A new emphasis on quality of care and patient safety was implemented, and the agency's name was changed to the Agency for Healthcare Research and Quality (AHRQ). Four years after its "near-death experience," AHRQ had regained all the funding it lost, and by 2002 the budget had grown to more than double its pre-1995 level.²¹ Wennberg has argued for an expanded role for AHRQ, noting that outcomes research has the potential to restrain wasteful spending and could help to control costs.¹³

In fact, the federal government is increasingly interested in supporting comparative effectiveness research to evaluate, for example, the efficacy of competing drugs, or to compare the effectiveness of different treatment options. In 2009 it announced plans to provide \$1.1 billion

to the AHRQ, the National Institutes of Health, and the Department of Health and Human Services to conduct the research, and it also provided funds to the Institute of Medicine to recommend priorities for spending the money.²²

As for treatment of low back pain, surgery rates in the Medicare population increased by 220 percent between 1988 and 2001, and the rates vary dramatically across geographic areas.²³ To determine what an appropriate rate might be, a prospective study was conducted in Maine, where surgery rates were four times higher in some areas than in others. The researchers followed all patients who had surgery to see whether their symptoms improved after the operation. They found that the best outcomes occurred in the areas where the rates were lowest; and the worst outcomes occurred in the areas with the highest rates. The evidence suggested that surgeons in the low-use area used more stringent criteria for recommending surgery. In these areas, patients with more severe disease were more likely to benefit, and those with less severe disease avoided the risks of surgery, which are significant. The authors concluded: "Outcomes research has the potential to provide information that will enable each patient to better understand the outcomes, risks and benefits of an operation and other treatment."²⁴(p.761)

Quality

The AHCPR drama came at a time when there had been a series of highly publicized medical errors. A 39-year-old health reporter for *The Boston Globe* died after receiving an overdose of a chemotherapy drug while being treated for breast cancer at one of the most prestigious hospitals in the country. A 51-year-old diabetic man had the wrong leg amputated in a Florida hospital. And an 8-year-old boy in another Florida hospital died due to a drug mix-up during "minor" surgery.

A number of studies were published in the 1990s documenting that preventable medical errors occurred in 1.5 to 2 percent of hospitalizations, and that many of these errors caused the patient's death. The Institute of Medicine was asked to investigate the issue and recommend a strategy that would lead to improvements in quality of care. The study led to the publication in 1999 of a report, *To Err Is Human: Building a Safer Health System*.²⁵ The report estimated that 44,000 to 98,000 deaths per year in the United States were caused by medical errors, more than motor vehicle accidents, breast cancer, or AIDS, placing medical errors among the top ten causes of death.

Before the Institute of Medicine report was published, medical errors were blamed on failures by individual doctors and nurses; practitioners who made mistakes were sued for malpractice, and some had even been prosecuted as criminals. The report shifted the blame to the medical care system—or nonsystem, according to some critics—characterizing it as decentralized and fragmented, rife with confusion, miscommunication, and lack of incentives

for improvements in safety. The Institute of Medicine committee compared the medical care industry unfavorably with other high-risk industries that had been much more successful at improving safety and preventing injury, especially the commercial airline industry. The report made a number of recommendations, beginning with the creation of a Center for Patient Safety within the AHRQ, which would set national goals, track progress, develop a research agenda, evaluate methods for identifying and preventing errors, and disseminate information. Another recommendation was that, as in the airline industry, accidents and near-misses should be reported so that errors could be investigated, leading to an understanding of the underlying factors that contribute to them. A mandatory, nonpunitive system should be developed that encourages providers to learn from their mistakes.²⁵

Recognizing that many adverse events involve medication errors, the report recommended that the Food and Drug Administration (FDA) should require that drug naming, packaging, and labeling be designed to minimize confusion. Because of doctors' notoriously poor handwriting, procedures should be developed to ensure accurate communication of prescriptions and other orders.

In 2009, Consumers Union (CU), the nonprofit agency that publishes *Consumer Reports*, published an evaluation of progress in implementing the Institute of Medicine report's recommendations ten years later.²⁶ The report gave the country a failing grade in implementing procedures they believe necessary to create a healthcare system free of preventable medical harm. In particular, CU reports that few hospitals have adopted measures to prevent medication errors and that the FDA rarely intervenes. Computerized prescribing and dispensing systems have not been widely adopted, despite evidence that they make patients safer. There is no national system of reporting medical errors and, where there is reporting, it is generally confidential, meaning that patients do not have access to information on how to compare the performance of doctors and hospitals, and there is little pressure for them to improve. Another Institute of Medicine recommendation was to raise standards for competency of doctors, nurses, and other healthcare professionals by requiring them to periodically pass examinations demonstrating skills, knowledge, and use of best-practice care in order to maintain their certification. Most specialty boards now have this requirement but, according to the CU report, there is no mechanism in place to ensure the competency of the 15 percent of physicians not certified by one of these boards, as well as those "grandfathered" prior to the adoption of the standards.²⁶

The CU report, as an example of medication errors, described the widely publicized incident in which the twin babies of actor Dennis Quaid and his wife were given one thousand times the prescribed dose of the blood thinner heparin because the different doses were packaged in similar vials with similar blue labels. The twins survived, but even though a similar mix-up had caused the deaths of three infants the previous year in an Indianapolis hospital, the packaging had not been changed.

An example of a system that works, described in the CU report, was instituted in 2004 in 103 Michigan intensive care units to prevent catheter-associated bloodstream infections. It consisted of a short checklist of best practices related to catheter use; nurses were empowered to ensure that doctors were following these practices. Researchers tracked catheter-associated infections and found that the incidence dropped to less than 20 percent of what it had been before the procedures were implemented. The Centers for Disease Control and Prevention has estimated that hospital-acquired infections kill 99,000 people each year.

The CU report argues that among the most important of the Institute of Medicine recommendations is "increased accountability through mandatory, validated and public reporting of preventable medical harm, including healthcare-acquired infections." According to the report, "It is a fundamental principle of quality control that if a process cannot be measured, it cannot be improved."^{26(p.6)}

Medical Care Report Cards

The rise of managed care (see Chapter 26) contributed to an increasing interest in the measurement of the quality and efficiency, or cost-effectiveness, of medical care. Managed care's focus on cutting costs, however, conflicted with the common assumption that, when it comes to medical care, more is better—an assumption that is challenged by outcomes research that suggests that sometimes less may be better as well as less expensive.²⁷ However, many people are suspicious that managed care companies, which have a financial incentive to do less for their patients, may have an inherent conflict of interest. The suspicion is especially strong in the case of for-profit managed care plans, which have an obligation to maximize profits for their investors, perhaps at the expense of the patients.

In the medical care marketplace, where economic factors are becoming increasingly significant, outcomes research has an important role to play in evaluating the quality and efficiency of different medical plans. In theory, when given enough information, customers—both the employers who choose which plans to offer and the employees who must choose among the plans that are offered—can make informed decisions, weighing quality and cost.²⁸ Moreover, patients are increasingly becoming more active participants in their own care. In part because of growing distrust of the medical system, patients want information on risks and benefits of available treatments and, if possible, on the competence of their physicians and other medical providers. Outcomes research provides some of this information.

Although managed care is often regarded with skepticism, it is more easily evaluated than the traditional fee-for-service form of medical practice. The organization of services that allows care to be "managed" also makes it possible for those services to be assessed in a formal way, something that is not realistic when each medical provider acts independently. Through an

accreditation process conducted by the nonprofit National Committee for Quality Assurance (NCQA), it is possible to rate managed care plans on their performance with respect to a number of standards. Information on the accreditation status of a plan can influence a business's decision about whether to offer the plan to its employees, and the information can be used by employees to choose among plans offered. In its 2008 State of Health Care Quality report, 845 health plans covering more than 100 million Americans provided data to NCQA on 54 different measures of healthcare quality. NCQA reported that most commercial plans had improved on most of the measures; Medicare and Medicaid beneficiaries did less well. There was significant geographical variation in how well the plans had measured up, with those in New England doing best and plans in South Central states doing worse than average. In general, health maintenance organizations (HMOs) did better than preferred provider organizations (PPOs).²⁹ Consumers can access "report cards" of plans on the NCQA website and compare their performances.

Many of the most easily measured standards used by NCQA focus on preventive care: for example, whether children receive a full set of immunizations and whether women get mammograms and Pap tests. Other standards evaluate how a plan manages care for patients with common diseases. The findings of outcomes research can be used, for example, to measure performance of an HMO in treating elderly heart attack victims. Research supported by AHCPR found that patients 65 years of age and older were 43 percent less likely to die after a heart attack if they were treated with beta blockers than if they did not receive these drugs.³⁰ Using that information, NCQA established, as one of its standards for evaluating a plan, the use of beta blockers for treatment of heart attacks. Since the agency began reporting on this measure, the percentage of heart attack patients who received the drugs went from 60 percent to well over 90 percent.³¹

Outcomes research can also be used in some circumstances to evaluate the performance of individual medical providers. The findings offer a basis not only for patients to choose where to go for treatment, but also for providers to compare their performance with that of their peers. Since 1989, New York State has measured the outcomes of coronary artery bypass surgery for treatment of blocked arteries in the heart, monitoring each of the 31 hospitals where the operations are performed. Mortality rates in 1989, adjusted for patients' risk factors such as age, diabetes, and hypertension, ranged between 0.88 percent and 10.02 percent.³² Data have also been collected on outcomes achieved by individual surgeons.

One of the study's findings was that hospitals that perform large volumes of coronary surgery have better outcomes than those that perform few of the operations, a result that has also been found true of other types of surgery. The New York study also found that surgeons who perform more than 150 bypass operations per year have only half the patient mortality rate of surgeons who perform fewer than 50. The publicity that followed release of the 1989 data on

individual hospitals led to a dramatic decline (41 percent) statewide in mortality rates associated with the surgery over the next three years.³³ Data collection and analysis continues, and mortality rates have continued to fall. By 1992, New York had the lowest risk-adjusted mortality rates of any state in the nation, and New York's rates continued to be lower than other states', at least through 1999. Thus, the information provided by outcomes research led to improved quality of surgical care statewide. An analysis of how the improvements were accomplished show that hospitals identified as performing poorly reacted strongly, for example, by restricting the surgical privileges of some low-volume surgeons whose patients were more likely to die from the operation.³⁴

Despite the successes, health services research has a long way to go before it can be widely used to help people make decisions about health care based on quality. Most of the indicators of managed care quality measured by accrediting agencies focus on preventive care for the healthy. Although this approach is important from a public health perspective, what matters most to patients is the quality of care they receive when they are ill.²⁷ Detailed analyses of providers' performance are available for only a limited number of procedures in New York, and most states would probably not be willing or able to carry out such ambitious programs. Surprisingly, although the New York State Health Department publishes annual reports on its cardiac surgery data (available at www.health.state.ny.us/statistics/diseases/cardiovascular), it appears that neither managed care companies nor patients use the information to choose hospitals and surgeons.³⁴

Inequities in Medical Care

Health services research has shed light on an unpleasant reality that pervades the American medical care system. Not only is care rationed by ability to pay, as discussed in Chapter 26, but there are racial inequities in how care is delivered even when individuals are able to pay for it. As documented in a 2002 Institute of Medicine report, *Unequal Treatment: What Healthcare Providers Need to Know about Racial and Ethnic Disparities in Healthcare*,³⁵ African Americans and Hispanics are less likely than whites to receive the most effective treatments for heart disease, human immunodeficiency virus (HIV) infection, asthma, breast cancer, and many other conditions, even when their income and insurance status are equal to whites.

Regarding heart disease, for example, the work of the New York State researchers described above has also found racial differences in access to coronary artery bypass surgery. It seems that physicians are less likely to recommend surgery to patients from ethnic minority groups than to comparable white patients. Studying files of patients who had undergone diagnostic testing in eight New York hospitals, and using guidelines developed by the RAND Corporation for "appropriateness" and "necessity" of the operation, the researchers selected 1261 patients who

would benefit from a coronary artery bypass. Returning to the files three months later, the researchers found that African-American and Hispanic patients were significantly less likely to have had the surgery than comparable white patients. It was not that the blacks and Hispanics had decided against the surgery; for the overwhelming majority, their physicians had not recommended it.³⁶

Childhood asthma is a chronic disease that can usually be kept under control by providing patients and their families with prescriptions for inhaled medications and education on how to use them. A study that examined records of young children hospitalized for asthma found that racial minorities were less likely than whites to have taken the most effective medications before they were hospitalized and were less likely to be given prescriptions for such medications when they were discharged. Thus, African-American and Hispanic patients received poorer quality care than whites, an observation that was especially disturbing because the prevalence of asthma in minority children is higher than in whites—25 percent higher.³⁷

According to the American Cancer Society, African Americans have the highest death rate and the shortest survival of any racial and ethnic group in the United States for most cancers. Although the overall racial disparity in cancer death rates is decreasing, the death rate for all cancers combined is 33 percent higher in black men and 16 percent higher in black women than in white men and women, respectively.³⁸ Blacks are less likely to survive five years after diagnosis, most likely due to a later stage at diagnosis, when the disease has spread. Blacks are also less likely to receive timely and high-quality treatment.

Blacks and Hispanics who are infected with HIV are less likely to receive antiretroviral therapy than nonminorities with HIV. Even after adjusting for insurance status, CD4 cell count, and other factors, minorities were 24 percent less likely than whites to receive protease inhibitors or other advanced drugs.³⁵

These studies provide evidence that inequities in medical care extend significantly beyond disparities in health insurance status and the corresponding financial barriers to care common in members of minority groups. The Institute of Medicine report concluded that “although myriad sources contribute to these disparities, some evidence suggests that bias, prejudice, and stereotyping on the part of healthcare providers may contribute to differences in care.”³⁵ However, more recent analyses indicate that the situation is more complex. Health services research by the Dartmouth group, discussed above, has found evidence that some of the differences are due more to geographic variations than racial disparities within the same area. Some of the disparities in treatment may be due to blacks living disproportionately in regions with low rates for all patients. Others may be due to higher-than-average surgery rates among whites rather than lower-than-average rates among blacks.³⁹

Other analysts remind us that the causes of disparities in health are not limited to disparities in health care. For example, there is a threefold difference in diabetes mortality rates between college graduates and those with only a high school education. No diabetes drug makes such a difference. It is clear that diabetes would be less of a problem in the United States if our society promoted education reform as avidly as it emphasizes healthcare reform.⁴⁰

The Relative Importance of Medical Care for Public Health

Health services research, in addition to studying medical care epidemiology, has tried to answer questions about the proper place for medical care in the public health system. To what extent does medical care contribute to improving the health of the population as a whole? Some skeptics have argued that medicine's effectiveness is limited and that its impact on health is marginal at best. As discussed in Chapter 1, much of the improvement in life expectancy over the past century resulted more from public health measures and improvements in the population's economic status than from improvements in medical interventions.

In focusing on the population perspective, analysts weigh the contribution of medical care with other factors that contribute to people's health (see **Figure 27-1**). There is little agreement on the relative importance of the various factors, which include genetics, lifestyle, and the environment, in addition to medical care. However, any consideration of these factors calls attention to the fact that, in the United States, resources devoted to medical care are far out of proportion to its contribution to health. In fact, the enormous American investment in medical care uses up resources that would otherwise be available to address other factors that affect health, such as education, housing, and the environment. In that sense, it may be that the greater the expansion of the medical care system, the more negative the impact on the population's health.⁴¹

Evidence from small-area comparisons in the United States, as well as comparative studies of industrialized nations, has clearly indicated that health is not correlated with resources devoted to medical care. This was true, for example, in Wennberg's comparison of healthcare costs and the population's health status in Boston and New Haven. Similarly, international studies of mortality rates in developed nations have found no consistent relationship with levels of medical care resources.⁴² The United States has higher rates of chronic disease prevalence and mortality than other OECD countries, despite its high spending on medical care. The fact that more medical care does not lead to better health is supported by a 2003 study by Wennberg's group that looked at patients with heart attacks, hip fractures, or colorectal cancer who lived in geographical areas with high Medicare spending compared with similar patients

in areas with lower spending. The researchers found that patients in the high-spending areas had more physicians' visits, more tests and procedures, and spent more time in the hospital than those in the low-spending areas, but the outcomes were not better and, in fact, included a small increase in the risk of death. Apparently, the higher-intensity practice patterns caused harm to patients.⁴³

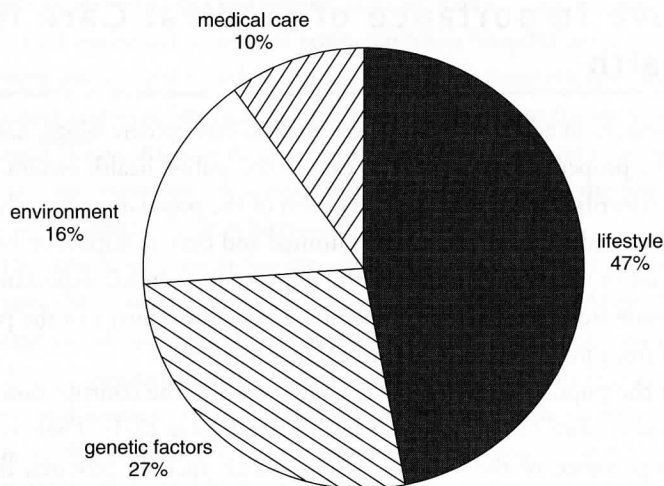


FIGURE 27-1 Contribution of Medical Care to Health. *Source:* Data from P.R. Lee, C.L. Estes, editors. *The Nation's Health* (Sudbury, MA: Jones & Bartlett Pub., 2001) p. 50.

As discussed in Chapter 26, the United States does not get its money's worth for the resources allocated to medical care. Health services research that focuses on the efficiency of the healthcare system can offer evidence on how the nation could keep costs under control while achieving better health. A number of studies have investigated the effects of different methods of paying for care on the use of care and on health outcomes. One influential study was the RAND Health Insurance Experiment.⁴⁴ This study compared use of services, expenditures, and health outcomes among several groups of consumers who were assigned randomly to receive free care or to pay copayments of varying amounts. The evidence showed, not surprisingly, that higher copayments discouraged patients from seeking care. The more consumers had to pay, the less medical care they consumed, and the free-care group used services costing 50 percent more than those who had to pay the most. For most of the participants, the extra services were not found to have any impact on their health status. Thus, many healthcare services provided to Americans with generous insurance policies may be wasted in that they do not contribute to better health.

However, for those who were poor and chronically ill, free care did provide significant benefits in health status. These are the people most likely to lack access to medical care because of financial barriers. They are the ones who may need care the most, but may be least likely to get it. This is the tragedy of the American health system which, despite the highest rate of healthcare spending in the world—much of it probably unnecessary—leaves 16 percent of the population uninsured.

At a time when the president and Congress are attempting to reform the American medical care system, there is a great deal to be learned from health services research. The research has demonstrated that a significant proportion of the resources spent on medical care in the United States does not contribute to better health in the population. If the nation can muster the political will to really reform how medical care is paid for in this country, the system could be made much more efficient and equitable than it is today, and the health of the American population could be significantly improved.

Conclusion

One hope for reducing costs of medical care and improving its quality is health services research, which studies the effectiveness, efficiency, and equity of the healthcare system. Small-area analysis, a form of health services research, has found that physicians in different geographical areas vary widely in how they treat common health problems. This observation suggests that for some conditions, decisions on treatment are somewhat arbitrary, and that different treatments may be equally valid—or invalid. Large variations are most likely when there is no clear evidence on which treatments are most effective.

The observed differences may be due in part to the varying availability of services in a community. Larger amounts of surgery are done in communities with higher numbers of surgeons. More people are hospitalized in communities with higher per capita numbers of hospital beds. Other variations seem to be merely variations in practice style, which tends to be shared by all physicians in a community. Comparisons of high-usage areas with low-usage areas have not found significant differences in health status, indicating that the variations are not caused by greater severity of illness in some areas, and there is no evidence that high usage helps or harms people's health. However, medical costs are proportionately high in the high-usage areas, suggesting that adopting the practices of low-usage areas could save substantial sums.

Outcomes research, the epidemiologic study of the everyday practice of medicine, holds hope for reaching conclusions on the benefits and risks of various treatments for high-variation conditions. For example, prostatectomy for benign prostatic disease is a common surgery performed on older men. However, it is not always effective in relieving symptoms, and it can have undesirable complications such as impotence and incontinence. Results of outcomes research

have made it clear that people should be informed of the risks of surgery and the possible outcomes before they make choices about their treatment. Surgery for BPH is now used less often than in the past. PSA screening for prostate cancer is also of questionable value since it leads to many biopsies and prostatectomies but does not appear to lower mortality rates.

The federal agency formerly called AHCPR got into political trouble in the 1990s when it published evidence recommending less use of back surgery for low back pain. Now known as AHRQ, the agency has regained its funding and more, and healthcare reformers are hoping that its comparative effectiveness research will help save money for the American healthcare system.

Outcomes research can be used to evaluate the quality of managed care plans, assessing whether plans provide services that have been demonstrated to be effective. The research can also be used to compare the performance of hospitals and surgeons. New York State has done this research, an exercise that has resulted in improved quality of coronary surgery in that state.

Health services research has documented extensive evidence that the delivery of medical care is inequitable and that ethnic and racial minorities receive poorer quality care than do white Americans. This is true even after differences in health insurance status have been taken into consideration.

Although health services research has been proven capable of improving the quality of medical care, it also shows that medical care is a less important influence on people's health than some other factors, including education, housing, and the environment. In fact, health services research suggests that if the United States spent less on medical care, and instead invested the savings in these other services, the population's health might be improved.

References

1. J. E. Wennberg, "Dealing with Medical Practice Variations: A Proposal for Action," *Health Affairs* 3 (1984): 6–32.
2. J. E. Wennberg and M. M. Cooper, eds. *Dartmouth Atlas of Health Care 1996* (Chicago, IL: American Hospital Publishing, 1996).
3. M. R. Chassin et al., "Does Inappropriate Use Explain Geographic Variations in the Use of Health Care Services? A Study of Three Procedures," *Journal of the American Medical Association* 258 (1987): 2533–2537.
4. E. A. Halm et al., "Has Evidence Changed Practice? Appropriateness of Carotid Endarterectomy After the Clinical Trials," *Neurology* 68 (2007): 187–194.
5. E. L. Hannan, C. Wu, and M. R. Chassin, "Differences in Per Capita Rates of Revascularization and in Choice of Revascularization Procedure for Eleven States," *BMS Health Services Research* 6 (2008): 35.
6. J. E. Wennberg, "Small Area Analysis and the Medical Care Outcome Problem," in L. Sechrest, E. Perrin, and J. Bunker, *Conference Proceedings: Research Methodology: Strengthening Causal Interpretations of Nonexperimental Data, AHCPR*. May 1990, 177–201.

7. G. Kolata, "In the U.S., All Medicine is Local," *The New York Times*, February 4, 1996.
8. J. E. Wennberg and M. M. Cooper, eds. *Dartmouth Atlas of Health Care 1998* (Chicago, IL: American Hospital Publishing, 1998).
9. J. E. Wennberg, "Outcomes Research, Cost Containment, and the Fear of Health Care Rationing," *New England Journal of Medicine* 323 (1990): 1202–1204.
10. M. Waldholtz and W. Bogdanich, "Warm Bodies: Doctor-Owned Labs Earn Lavish Profits in a Captive Market," *Wall Street Journal*, March 1, 1989.
11. B. J. Hillman et al., "Frequency and Costs of Diagnostic Imaging in Office Practice—A Comparison of Self-Referring and Radiologist-Referring Physicians," *New England Journal of Medicine* 323 (1990): 1604–1608.
12. R. Abelson and M. Petersen, "An Operation to Ease Back Pain Bolsters the Bottom Line, Too," *The New York Times*, December 31, 2003.
13. F. Mullan, "Wrestling With Variation: An Interview with Jack Wennberg," *Health Affairs Web Exclusive* (2004): VAR-71-80.
14. F. J. Fowler et al., "Symptom Status and Quality of Life Following Prostatectomy," *Journal of the American Medical Association* 259 (1988): 3018–3022.
15. J. H. Wasson et al., "Transurethral Resection of the Prostate among Medicare Beneficiaries: 1984–1997. For the Patient Outcomes Research Team for Prostatic Diseases," *Journal of Urology* 164 (2000): 1212–1215.
16. G. Lu-Yao et al., "Natural Experiment Examining Impact of Aggressive Screening and Treatment on Prostate Cancer Mortality in Two Fixed Cohorts from Seattle Area and Connecticut," *British Medical Journal* 325 (2002): 740–746.
17. M. J. Barry, "Screening for Prostate Cancer—The Controversy That Refuses to Die," *New England Journal of Medicine* 360 (2009): 1351–1354.
18. J. Kosterlitz, "Cookbook Medicine," *National Journal* (March 9, 1991): 574–577.
19. Agency for Health Care Policy and Research, "AHCPR Releases Low Back Pain Guideline" *Research Activities* (January 1995): 15–16.
20. N. A. Lewis, "Agency Facing Revolt After Report: Enraged Back Surgeons Recruiting Republicans for a Battle," *The New York Times*, September 14, 1995.
21. B. H. Gray, M. K. Gusmano, and S. R. Collins, "AHCPR and the Changing Politics of Health Services Research," *Health Affairs* (June 25, 2003), Web exclusive.
22. M. Mitka, "Studies Comparing Treatments Ramp Up," *Journal of the American Medical Association* 301 (2009): 1975.
23. R. A. Deyo et al., "Overtreating Chronic Back Pain: Time to Back Off?" *Journal of the American Board of Family Medicine* 22 (2009): 62–68.
24. R. B. Keller et al., "Relationship between Rates and Outcomes of Operative Treatment for Lumbar Disc Herniation and Spinal Stenosis," *Journal of Bone and Joint Surgery* 81-A (1999): 752–762.
25. Institute of Medicine, *To Err Is Human: Building a Safer Health System* (Washington, DC: National Academy Press, 1999).
26. Consumers Union, *To Err Is Human—To Delay Is Deadly*, www.safepatientproject.org/pdf/safepatientproject.org-to_delay_is_deadly-2009_05.pdf, May 2009, accessed February 23, 2010.
27. C. R. Gaus and L. Simpson, "Reinventing Health Services Research," *Inquiry* 32 (1995): 130–133.

28. M. Angell and J. P. Kassirer, "Quality and the Medical Marketplace—Following Elephants," *New England Journal of Medicine* 335 (1996): 883–885.
29. National Committee for Quality Assurance, News Release: "Report: Health Care Quality Improves but Varies across Different Regions of the Country," www.ncqa.org/tabid/847/Default.aspx, October 2, 2009, accessed September 27, 2009.
30. S. B. Soumerai et al., "Adverse Outcomes of Underuse of Beta-Blockers in Elderly Survivors of Acute Myocardial Infarction," *Journal of the American Medical Association* 277 (1997): 115–121.
31. National Committee for Quality Assurance, "The Basics: Performance Measurements," www.ncqa.org/tabid/441/Default.aspx, 2009, accessed September 27, 2009.
32. E. L. Hannan et al., "Improving the Outcomes of Coronary Bypass Surgery in New York State," *Journal of the American Medical Association* 271 (1994): 761–766.
33. E. L. Hannan et al., "Decline in Coronary Artery Bypass Graft Surgery Mortality in New York State: The Role of Surgeon Volume," *Journal of the American Medical Association* 273 (1995): 209–213.
34. M. R. Chassin, "Achieving and Sustaining Improved Quality: Lessons from New York State and Cardiac Surgery," *Health Affairs* (July/August 2002): 40–51.
35. Institute of Medicine, *Unequal Treatment: What Healthcare Providers Need to Know about Racial and Ethnic Disparities in Healthcare* (Washington, DC: National Academy Press, 2002).
36. E. L. Hannan et al., "Access to Coronary Artery Bypass Surgery by Race/Ethnicity and Gender Among Patients Who Are Appropriate for Surgery," *Medical Care* 37 (1999): 68–77.
37. J. A. Finkelstein et al., "Quality of Care for Preschool Children with Asthma: The Role of Social Factors and Practice Setting," *Pediatrics* 95 (1995): 389–394.
38. American Cancer Society, "Cancer Facts and Figures for African Americans 2009–2010," www.cancer.org/downloads/STT/cffaa_2009-2010.pdf, accessed September 30, 2009.
39. K. Baicker et al., "Who You Are and Where You Live: How Race and Geography Affect the Treatment of Medicare Beneficiaries," *Health Affairs* (October 7, 2004), Web exclusive.
40. S. H. Woolf, "Social Policy as Health Policy," *Journal of the American Medical Association* 301 (2009): 1166–1169.
41. R. G. Evans and G. L. Stoddart, "Producing Health, Consuming Health Care," *Social Science and Medicine* 31 (1990): 1347–1363.
42. G. F. Anderson, B. K. Frogner, and U. E. Reinhardt, "Health Spending in OECD Countries in 2004: An Update," *Health Affairs* 26 (2007): 1481–1489.
43. E. S. Fisher et al., "The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care," *Annals of Internal Medicine* 138 (2003): 288–298.
44. L. A. Aday, *Evaluating the Medical Care System: Effectiveness, Efficiency, and Equity* (Ann Arbor, MI: Health Administration Press, 1993).