VARIATION IN USE OF MEDICARE SERVICES AMONG REGIONS AND SELECTED ACADEMIC MEDICAL CENTERS: IS MORE BETTER?

John E. Wennberg
Dartmouth Medical School

December 2005

ABSTRACT: Initiatives for improving the quality of health care are now focused on stemming the underuse of "effective care"—therapy viewed as medically necessary care on the basis of clinical outcome evidence. But only a small proportion of the health care dollar is influenced by effective care. Most of the spending, at least regarding Medicare, is in two other categories. "Preference-sensitive care," in which treatment options involve tradeoffs that should be based on the patient's own values, tends not to be underused but misused. And "supply-sensitive care," in which the supply of resources governs the frequency of their use, is overused, particularly in the management of chronic illness. Hospital-specific measures that profile performance—such as the average number of days spent in the hospital during the last six months of life and physician labor inputs over that time—could help identify more efficient providers.

This report is based on the annual Duncan W. Clark Lecture delivered by the author at the New York Academy of Medicine on January 24, 2005.

CONTENTS

LIST OF FIGURES

ABOUT THE AUTHOR

John E. Wennberg, M.D., M.P.H., is director of the Center for the Evaluative
Clinical Sciences at Dartmouth Medical School. He has been a professor in the
Department of Community and Family Medicine since 1980 and in the Department of
Medicine since 1989, and currently holds the Peggy Y. Thomson Chair for the Evaluative
Clinical Sciences. With colleague Alan Gittelsohn, he developed a method of determining
population-based rates for the utilization and distribution of health care services. This
method, called small area analysis and first published in 1973, revealed large variations in
health-care usage among different areas. Work to uncover the reasons behind these
variations led Wennberg and his colleagues to develop techniques to document the results
of common medical practices, a strategy that came to be called outcomes research. In
addition to his work on the *Dartmouth Atlas of Health Care*, he is also exploring the use of
interactive video technology to inform patients of the results of outcomes research so they
can participate in medical decision-making. Wennberg received his medical degree from
McGill University Faculty of Medicine and his master of public health degree from Johns
Hopkins University Bloomberg School of Public Health.

EXECUTIVE SUMMARY

Initiatives for improving the quality of health care are now focused on stemming the underuse of "effective care"—therapy that is viewed as medically necessary care on the basis of clinical-outcome evidence, preferably from randomized trials. An example is the use of a beta-blocker drug after a heart attack. Causes of such underuse include discontinuity of care (worsened when too many physicians are involved) and lack of infrastructure to assure outreach and the timely use of effective-care services. Pay-for-performance strategies should reduce such underuse.

But while giving providers incentives to do the things they ought to do will very likely increase the use and quality of effective care and save lives, it is unlikely to have a major impact on rising costs; only a relatively small proportion of the health care dollar is influenced by effective care. Most of the spending, at least regarding Medicare, is in other categories—"preference-sensitive care" and "supply-sensitive care"—in which the quality problem is not underuse.

Preference-sensitive care, in which treatment options involve significant tradeoffs that should be based on the patient's own values, tends not to be underused but misused. The causes of this misuse include failure to accurately communicate the risks and benefits of the alternative treatments and the failure to base choice of treatment on the patient's opinion rather than those of others. Adjustment of economic incentives to reward adopters of shared decision-making could lead to a reduction in such unwarranted variation.

The third category of care—supply-sensitive care, in which the supply of resources governs the frequency of their use—is overused, particularly in the management of chronic illness. The causes include overdependence on acute hospital care and lack of infrastructure to support continuous management of chronically ill patients in other care settings. Ironically, populations receiving more supply-sensitive care do not have better outcomes. In one study—in which researchers examined the outcomes of three sets of patients (who had either a hip fracture, heart attack, or colectomy for colon cancer) and followed them for up to five years—the major finding was that regions with greater care intensity showed increased mortality rates.

Hospital-specific measures that profile performance in managing chronic illness could help identify more efficient providers. Moreover, pay-for-performance strategies, along with related strategies to reward efficient providers and pay for chronic-illness-management infrastructure, could promote reform.

In that spirit, the author and his colleagues in the Dartmouth Atlas Project profiled the management styles of 77 hospitals, most of them well-known academic medical centers that had been rated by *U.S. News and World Report* as the nation's "best" for treating geriatric care, heart disease, cancer, and pulmonary disease. Concentrating on patients' last six months of life, the researchers gathered data on several measures: average number of days spent in the hospital during that time, average number of days spent in intensive care units, average number of physician visits, percent of patients who see 10 or more physicians, percent of patients who die in intensive-care units, Medicare spending, and physician labor inputs.

Although selected for their reputations for high-quality care, these hospitals differed remarkably amongst themselves in the way they managed severely ill Medicare patients. This was often true even among hospitals in the same state or city.

The Dartmouth Atlas Project recently made hospital-specific information available for California, and plans to do subsequent releases regarding other parts of the United States. The simple availability of information on the relative efficiency of specific health care organizations in managing chronic illness could prove beneficial. It may stimulate payers to reexamine their provider networks and motivate employers to steer their employees toward efficient hospitals.

In the long run, the most challenging problem will be finding mechanisms to clear regional markets of excess capacity. While special deals made with forward-thinking providers may well result in models of how to deliver care that is simultaneously of high quality and low cost, strategies to assure that all Medicare patients are served by such hospitals remain elusive. If Medicare administrators were willing and able, however, to take steps to select providers on the basis of quality and efficiency—and other payers were willing to play by similar rules—this would serve as a life-or-death wakeup call to the provider community, and it would likely result in accelerated change throughout the nation's health care markets.

VARIATION IN USE OF MEDICARE SERVICES AMONG REGIONS AND SELECTED ACADEMIC MEDICAL CENTERS: IS MORE BETTER?

INTRODUCTION

By some accounts, health care in the United States has entered a death spiral of ever-escalating costs and progressive loss of entitlement—more and more employers are electing not to provide health insurance, and those who do tend to shift the financial burden onto their employees. At the same time, Medicare appears headed toward fiscal ruin.

Some still hold out the hope that what has become known as "pay-for-performance" will save the day. Instead of applying the same rate to all providers, those whose practices show excellent performance in meeting high-quality-care guidelines would be rewarded with higher reimbursements. Others believe that the answer lies in making consumers better purchasers of health care through modifications of health insurance. Wiser spending through high deductibles and medical savings accounts, it is argued, would lead to a more rational medical market.

Our own studies of practice variations hold some good news and some bad news for both kinds of efforts.

Quality initiatives are now focused on stemming the underuse of "effective care"—therapy that is viewed as medically necessary care on the basis of clinical outcome evidence, preferably from randomized trials. An example is the use of a beta-blocker drug after a heart attack. But while giving providers incentives to do the things they ought to do will very likely improve the quality of care and save lives, it is unlikely to have a major impact on rising costs; only a relatively small proportion of the health care dollar is influenced by effective care. Most of the spending, at least regarding Medicare, is in other categories of care in which the quality problem is not underuse.

More than 50 percent of Medicare spending is used to buy "supply-sensitive" health care—visits to physicians, diagnostic tests, and hospitalizations, mostly for patients with chronic illnesses. Here the most important problem is *over*use—more is not necessarily better, particularly with regard to inpatient care. People with chronic illnesses who live in regions where both health care resources and health care spending are higher do not have better health outcomes. In fact, in some cases they have somewhat shorter life expectancies than people who live in regions where resources are less abundant and less inpatient care is used for the management of patients with chronic illnesses. Overuse thus

1

has two consequences: 1) the health care system spends more money without achieving a benefit; and 2) patients are exposed to the burdens and risks of treatment that is unnecessary or counterproductive.

As a tool for addressing the use of care among the chronically ill, the high-deductible health plan and medical savings account strategies are problematic. Because the volume and costs of such care become progressively higher as illness progresses—reaching a crescendo toward the end of life—even well-endowed savings accounts may soon be exhausted and thus have little influence.

Another significant portion of Medicare spending is for "preference-sensitive" care, epitomized by discretionary surgery. In this case, misuse of care is the problem, with use of medical services driven more by provider opinion than by informed patient preference. A pay-for-performance initiative that rewarded providers for encouraging patients to participate in informed decision-making might have the effect of decreasing demand for surgery (since informed patients generally choose less aggressive treatment strategies than what physicians prescribe for them). Such an incentive program might have some economic effect on demand, though its impact would be limited.

In cases where the goal is to increase appropriate utilization—such as in immunization and other examples of effective care that are currently underused—it is hard to see how financial considerations such as high deductibles, which discourage patient access, can help improve quality.

This report has three objectives. The first is to demonstrate that categorizing health care services into "effective care," "preference-sensitive care," and "supply-sensitive care" is a useful way to view unwarranted practice variations and to help devise initiatives that address them. The second is to review recent progress, using Medicare claims data, in developing provider-specific performance measures. Finally, the third objective is to briefly consider the requirements for achieving real and sustainable improvements in quality and efficiency in each of the three posited categories of care.

EFFECTIVE CARE

In the effective-care category, the benefits are thought to so outweigh the risks that virtually all patients with a specific medical need should receive the service. Most effective-care services, however, are underused. For a 2003 study published in the *New England Journal of Medicine,* Elizabeth McGlynn and her colleagues used a sample of medical records to examine compliance with practice guidelines, most of which

targeted the underuse of effective care. Data were obtained on 439 quality measures, and the researchers indeed found that patients received recommended care less than 55 percent of the time.[1]

The Dartmouth Atlas Project has had only limited success in measuring effective care using claims data, either because the population at need (e.g., the subgroup of heart attack patients needing beta-blockers at discharge) cannot be accurately defined in the claims, or the item of necessary care is not paid for by Medicare. Several services can be calculated, however; for those that could be measured, extensive underuse of effective care was found. For example, practice guidelines call for an eye examination at least once every two years for people with diabetes (Figure 1). Yet in several hospital-referral regions in 2001, fewer than 50 percent of Medicare enrollees with diabetes had eye examinations; even in the "best regions," only about 75 percent of enrollees had them. In locales in and around New York City, rates were above average but not exemplary. For example, 64 percent of diabetic residents of Manhattan received recommended care, and in the Bronx the rate was slightly lower, 63 percent.

The underuse of effective care relates in large part to the lack of the infrastructure necessary to support systematic compliance with guidelines. Thus, when organized group practices such as Kaiser Permanente have made concerted efforts to improve the management of chronic illness, including the development of processes that identify patients in need and ensure that the proper treatment is provided, these efforts have led to rates of guideline compliance greater than those of fee-for-service medicine. Similarly, enrollees in traditional Medicare in regions or states with fewer specialists and more family practice physicians (and less Medicare per capita spending) are more likely to receive effective care. By contrast, patients with chronic illnesses who live in high-spending regions tend to have many more physicians involved in their care, raising questions about who is in charge and responsible for ensuring that needed care is delivered.[2]

Identifying patients in need will become easier as electronic medical records become more widely used, and the adoption of such technology may be accelerated by pay-for-performance. However, because underuse of effective care is not associated with overall Medicare spending, one should not assume that doing the right thing here will lead to a reduction in per capita spending. *To have a significant impact on Medicare costs, pay-for-performance strategies must be directed not so much to effective care but to the other two categories— preference-sensitive care and supply-sensitive care.*

PREFERENCE-SENSITIVE CARE

Preference-sensitive care typically involves significant tradeoffs that affect the patient's quality or length of life. The surgical options for treating early stage breast cancer, for example, usually include mastectomy (complete removal of the breast) or lumpectomy (a local excision of the tumor), often called "breast-sparing surgery." The consequences for women who choose mastectomy include the loss of the breast and, for some, the use of a prosthesis or the undergoing of reconstructive surgery. For women who choose breast-sparing surgery, consequences can include radiation or chemotherapy, or both, and living with the risk of local recurrence, which would require further surgery.

The Dartmouth Atlas Project has noted striking regional variations in the proportion of early stage breast cancer patients who undergo lumpectomy. In an early study (1992–93), regions were identified in which virtually no Medicare women underwent lumpectomy, while in one region nearly 50 percent did. Even adjoining regions sometimes had strikingly different rates. For example, in the Elyria, Ohio, hospital-referral region, 48 percent of Medicare women had breast-sparing surgery for early-stage breast cancer, while Cleveland and Columbus registered only 23 percent and 12 percent, respectively.

Many of us believe that the major source of such widely varying discretionary surgery rates is idiosyncratic practice style. This theory was first advanced in the 1930s by J. Alison Glover, a British pediatrician, whose studies revealed a near tenfold variation in tonsillectomy rates among school districts. One of Glover's important findings was that the decision of whether or not to perform a tonsillectomy was made by a single physician— the school health officer who routinely examined students for signs of illness—and his most convincing evidence was the "natural experiment" that occurred with the arrival of a new health officer in the Hornsey Borough school district. Within a year, the rates of tonsillectomy in the district dropped by a factor of 10, and they remained low for years afterward. Glover attributed the contrasting rates to the change in "medical opinion" embodied in the different practice styles of the two physicians.

Similarly, the author, together with his colleague Alan Gittelsohn and two physicians from Morrisville, Vt., reported a tenfold variation in tonsillectomy rates among Vermont regions in the early 1970s. After the Morrisville physicians became aware of the high rate in their own area, local medical opinion changed radically and the town's rates dropped nearly to the bottom of the distribution.

A common rebuttal to the practice-style theory is that patient preferences actually dominate decision-making, and that rates of surgery are thus proportional to variations in preferences. Under this alternative theory, the interpretation would be that while 48 percent of Elyria women with early stage breast cancer preferred lumpectomy, only 12 percent in Columbus did and exceedingly few women in Rapid City, South Dakota— a mere 1 percent—did. These two theories might be a subject of legitimate debate if the physician's recommended course of treatment corresponded reasonably closely to the patient's informed preference. But experimental evidence from clinical trials of shared decision-making aided by patient decision aids shows that when it comes down to choosing treatment options, physicians' opinions and patients' preferences are *not* well correlated.

Shared decision-making is the process of interacting with patients to help them "make informed, values-based choices among two or more medically reasonable alternatives," and patient decision aids are "standardized, evidence-based tools designed to facilitate that process."[3] They are designed to provide: (1) high-quality, up-to-date information about the condition, including risks and benefits of available options and, if appropriate, a discussion of the limits of scientific knowledge about outcomes; (2) values clarification to help patients in sorting out their beliefs and preferences; and (3) guidance or coaching in deliberation so that the patient's involvement in decision-making may be improved.

Clinical trials of patient decision aids have now been completed for a number of conditions involving discretionary surgery. They include: the choice between lumpectomy and mastectomy for early stage breast cancer; the choice between invasive cardiac treatment or more conservative medical management for chest pain resulting from coronary artery disease; and the choice between surgery and conservative management for patients with back pain caused by disk disease. The trials show that, compared with a control group, patients who use decision aids are better informed about the benefits, risks, and clinical uncertainties associated with the treatment options available to them. Moreover, the choices patients make in the shared decision-making environment—when assisted by patient decision aids—are "better" decisions: they more closely reflect the patient's own individual values. Finally, most of these clinical trials show a net reduction in demand for the more invasive surgical options, an outcome of particular importance for the health care economy.

The last point deserves amplification. In "usual practice," where physicians presumably base their judgment on clinical evidence, the supply of patients in a given region whose level of illness makes them clinically appropriate candidates for surgical

intervention may well exceed the amount of surgery actually being done in that region.[4] A recent study by Hawker and colleagues—of arthritis patients deemed able to benefit from knee surgery, should it be performed—speaks to this point. The number of patients "in need" (defined as clinically appropriate for surgery) exceeded the rate of surgery for the corresponding age and sex groups by a factor of more than 10. The most important finding, however, was the striking contrast between need for surgery as defined by physicians and need as defined by patient preferences. When these patients were interviewed concerning their preference for treatment, only 14 percent indicated a preference for surgery; the vast majority wanted conservative treatment.

Such informed patient involvement, or the lack of it, produces wide differentials region by region in the frequency of invasive procedures. In examining, for example, the distribution in rates among hospital-referral regions of the three orthopedic procedures of knee replacement, hip replacement, and back surgery, it is seen that they all vary remarkably, particularly when compared to hip-fracture repair. Knee replacement and hip replacement are respectively four and five times more variable than hip-fracture repair, and back surgery is about seven times more variable (Figure 2).

The sometimes remarkable differences among neighboring regions is exemplified by the "surgical signatures" of four South Florida communities. Figure 3 compares the rates of surgery in Miami, Fort Lauderdale, Fort Myers, and Sarasota to rates in Manhattan (which serves as a base case because the rates there are among the lowest in the nation). This comparison might be of particular interest because Medicare residents of Manhattan commonly winter in Florida. In the years 2000 and 2001, the rate of knee surgery in Fort Myers was three times higher than that of Manhattan. The rate in Sarasota was 2.5 times higher, and the rate in Fort Lauderdale was 1.8 times higher. Among these same communities, the rates of hip replacement were twice the rate of Manhattan; and back-surgery rates were over three times higher in Fort Myers and Sarasota and two times higher in Fort Lauderdale. By contrast, the rates for Miami were much closer to those of Manhattan than to the other South Florida medical communities: Hip replacement rates were 11 percent lower in Miami while the rate of knee surgery was 26 percent higher and the rate of back surgery was 39 percent higher.

In theory, the variations among these communities in rates of knee replacement, hip replacement, and back surgery could reflect differences in patient preferences about treatment or the incidence patterns of osteoarthritis and herniated disks. In light of the evidence, this seems unlikely. Moreover, there is no epidemiologic evidence that illness rates or informed patient preferences vary as sharply from one health care market to

another as does surgery. It seems very unlikely, for example, that differences in illness incidence or patient preference could account for rates of knee, hip, and back surgery in Fort Myers being twice what they are in Miami, or for the peculiar distributions of orthopedic procedures that favor back surgery over knee replacement (as in Sarasota) or knee replacement over hip replacement (as in Fort Myers).

The behavioral basis of the surgical-signature phenomenon seems to lie in the propensity of local surgeons to specialize in a particular subset of the orthopedic surgical workload—they could, for example, choose trauma, sports medicine, or carpal tunnel syndrome, as well as knee, hip, or back conditions—and in their ability to find candidates that meet clinical appropriateness criteria. In Fort Myers, surgical workloads are oriented toward knee and back surgery; in Sarasota, back surgery is favored over knee and hip replacement; and in Fort Lauderdale, the rate of hip replacement is higher than the U.S. average.

An examination of the association between the per capita supply of surgical specialists and the rates of procedures that each specialty performs adds further insight. If surgeons of a particular specialty were allocating their time and surgical effort among a prioritized list of indications based on patients' needs and preferences, regions with more surgeons should have higher rates of surgery for common conditions such as osteoarthritis of the knee and hip. But in fact there is very little association between the supply of orthopedic surgeons and the rates of hip and knee surgery. For example, although the per capita supply of orthopedic surgeons varies more than 4.7-fold among regions, there is no relationship between the supply of orthopedic surgeons and rates of knee replacement, and there is little relationship with hip replacement. (The correlations between supply and surgery rates have R^2 values of .01 and .06, respectively, for knee and hip. That is, only 1 percent and 6 percent of the variations in surgery rates are "explained" by the supplies of associated surgeons. The relationship between the supply of orthopedic surgeons and rates of back surgery has an R^2 value of .02.[5])

The persistence of surgical signatures over long periods supports the interpretation already suggested: surgical specialists tend to become expert in a subset of the procedures that their specialty performs, and they orient their workload toward patients eligible for the procedure with which the surgeon is most comfortable. Figure 4 shows the surgical signatures of the Fort Myers, Fort Lauderdale, and Miami hospital-referral regions over a decade, as ratios of the local rates relative to the Manhattan rates. Note the year-in, year-out consistency in the rates. Note, too, that over the decade the differences in rates add up to substantial differences in the numbers of procedures performed. For example, the

surgeons working in Fort Myers performed 7,246 more back operations, 7,099 more knee replacements, and 2,689 more hip replacements than would have been done had the Manhattan rates prevailed in those communities.

The stability of the surgical signatures of orthopedic procedures in Fort Myers, Fort Lauderdale, and Miami is typical of the nation as a whole, as evidenced by the strong correlation between regional rates of a given procedure in 1992–93 and the rates in 2000–01. The R^2 correlation between knee replacement rates in those two periods is .75. Interestingly, while the United States average rate of these surgeries increased by 40 percent over those years, and the supply of orthopedic surgeons increased about 9 percent, local practice patterns changed little. Variations among regions simply don't show a strong tendency to "regress to the mean." Similar patterns were evident in hip replacement and back surgery, where the correlations between rates in 1992–1993 and 2000–2001 had R^2 values of .81 and .51, respectively.

Is More Better?

In the early 1990s, an opportunity presented itself for testing the assumption that the systematic implementation of shared decision-making supported by decision aids (free of undue influence on patients from the practice styles of their physicians or other inappropriate pressures) would produce the "right rate"—the actual demand—for a given treatment option. A decision aid designed to help patients decide between watchful waiting and surgery for their enlarged prostates was introduced in the urologic clinics of two prepaid group practices, Kaiser Permanente in Denver and Group Health Cooperative in Seattle. After the implementation of shared decision-making, the population-based rates of prostatectomy fell 40 percent, providing a measure of demand when patients are informed and involved in the choice of treatment. (Rates in the control group, Group Health Cooperative's Tacoma site, did not change.) The rate resulting from shared decision-making was actually at the extreme low end of the national distribution, suggesting that prostate surgery in most regions of the United States occurs substantially more often than informed patients would actually wish.[6]

One could extend this result by speculating that the amount of discretionary surgery, of all types, performed in the United States exceeds the amount that informed patients want. What is safe to conclude, however, is that current patterns of practice do not reflect demand based on patient preferences. Geographic variations in rates of surgery, which largely reflect physician practice style, will persist until patients are actively involved in the decision process and there are incentives for physicians to adopt such shared decision-making.

8

The introduction of shared decision-making for preference-sensitive care involving discretionary surgery could have significant economic impact on a health care market. For example, over the 10-year period 1992–2001, Medicare spending (in 2001 dollars) for knee and hip replacement and back surgery in Fort Lauderdale and Fort Myers is estimated to be, respectively, $137 million and $135 million more than would have been spent if the Manhattan rates had prevailed. In Miami, the excess spending amounted to $25 million. A change in utilization that more accurately reflected "true" patient-driven demand could result in cost savings for the payers and better quality of care for patients.

SUPPLY-SENSITIVE CARE

The third category of care, supply-sensitive care, reflects the generally held assumption that the supply of resources governs the frequency of their use, especially for people with chronic illnesses—notably, congestive heart failure, chronic lung disease, and cancer. The level of spending on these conditions reflects the frequency of physician visits (and revisits), hospitalizations, stays in intensive care units, referrals to specialists, and the use of imaging and other diagnostic tests.

Overall, supply-sensitive care appears to account for some 50 percent of medical spending, though there is remarkable variation in the frequency of use of these services among regions. For example, rates of primary care visits vary by a factor of about three, visits to medical specialists by more than six, and hospitalizations for cancer, chronic lung disease, and congestive heart failure by more than four (Figure 5). The use of hospitals for the treatment of people with medical conditions is particularly intense during the last few months of life, and its variation among regions is striking. On average, patients living in the lowest-rate regions spend about six days in hospitals while those in the highest-rate region spend 20 days (Figure 6).

In contrast to effective care and preference-sensitive care, where clinicians have strong opinions on the need for specific interventions, medical theories and medical evidence play little role in governing the frequency of use of supply-sensitive services.[7] For patients at a given stage in the progression of chronic illness, medical textbooks contain no evidence-based clinical guidelines for scheduling them for return visits, when to hospitalize or admit them to intensive care, when to refer them to a medical specialist, and, for most conditions, when to order a diagnostic or imaging test. As an example, the *British Medical Journal*'s annual *Clinical Evidence Concise*—which describes itself as "the international source of the best available medical evidence for effective health care"—contains not a single reference as to when to hospitalize, or schedule for a revisit, patients with cancer, chronic lung disease, or heart failure.[8]

Demand for such services has instead been driven by their supply. For example, the Dartmouth Atlas Project has consistently shown over the years a positive association between the supply of staffed hospital beds per 1,000 residents and the hospitalization rate for medical (nonsurgical) conditions (Figure 7). The effect of hospital bed supply on hospital use is so well recognized, in fact, that it has often been referred to as "Roemer's law."[9] There are exceptions, however. Hospitalizations for hip fracture—one of the few conditions in which variation closely reflects the incidence of illness—correlate little with resource supply. And hospitalizations for major surgery, whether in the preference-sensitive or effective care categories, are not correlated with overall beds per capita.

A similar relationship can be seen between the supply of physicians and visit rates, particularly for those specialties focused on treating chronic illnesses. For example, in Figure 8 about half of the variation in the number of visits to cardiologists in the region per Medicare enrollee is associated with the number of cardiologists per 100,000 residents. Such a relationship makes arithmetic sense: on average, regions with twice as many cardiologists per 100,000 residents will have twice as many available office visit hours, especially as appointments to see physicians characteristically are fully "booked"—very few hours in the work week go unfilled. In the absence of evidence-based guidelines on the appropriate interval between visits, available capacity governs the frequency. A similar relationship exists between the supply of internists and numbers of visits to internists.

Physician visit rates among people who are in their last six months of life vary substantially as well. In the highest-rate region, terminal patients had an average of more than 55 visits during their last six months; in the lowest-rate regions the average was about 14 visits (Figure 9).

Is More Better?
The bottom line question is whether populations receiving more supply-sensitive care have better outcomes. Do they live longer? Do they have higher quality of life? Are they more satisfied with their care? As might be deduced from the absence of practice guidelines, this issue has received virtually no attention from academic medicine or from federal agencies, such as the National Institutes of Health, that are responsible for the scientific basis of medicine. With the exception of a few studies of chronic disease management, patient-level studies that might shed light on the question simply have not been done. The appropriate quantity of supply-sensitive care is only now beginning to emerge, at medical rounds and in scientific journals and textbooks, as a topic for medical discourse.

A recent population-level study by Elliott Fisher and colleagues at Dartmouth provides a provisional answer about whether regions with greater intensity of clinical

practice have better outcomes.[10] The researchers examined the outcomes of three patient cohorts enrolled because they had either a hip fracture, heart attack, or colectomy for colon cancer, and the patients were followed for up to five years after their initial event. The study's major finding: regions with greater care intensity showed increased mortality rates.

Figure 10, adapted from the Fisher study, compares the level of resource inputs and mortality among cohorts living in two regions—the highest and lowest quintiles in Medicare end-of-life spending. The high-rate regions had 32 percent more hospital beds per capita, 31 percent more physicians, 65 percent more medical specialists, 75 percent more general internists, 37 percent more surgeons—and, of course, more Medicare spending (61 percent higher, on a price-adjusted basis). The low-rate regions, for their part, had 25 percent more family practice physicians.

Although the hip fracture, colon cancer, and heart attack cohorts were comparable in baseline morbidity, those living in the high-rate regions had higher mortality rates: 1.9 percent higher for hip fracture patients, 1.2 percent higher for colon cancer patients, and 5.2 percent higher for heart attack patients.

What about functional status and patient satisfaction? To address this question, Fisher and colleagues used a fourth data set, the ongoing Medicare Current Beneficiary Survey, which contains measures of functional status and patient satisfaction. The results indicated no difference between regions in functional status or satisfaction, but lowered access to patient care in high-rate regions.

Fisher and colleagues repeated their study of regional outcomes, this time restricting the study to focus on patients who received their initial care at academic medical centers. The results were quite similar: academic medical centers in high-intensity regions provided more supply-sensitive services than those in low-intensity regions. For example, during the first six months following hip fracture, patients using academic medical centers in high-spending areas had 82 percent more physician visits, 26 percent more imaging exams, 90 percent more diagnostic tests, and 46 percent more minor surgery. Nevertheless, patients in high-intensity regions had higher mortality rates and worse "score cards" on measures of quality.

HOW WELL-KNOWN ACADEMIC MEDICAL CENTERS MANAGE SEVERE CHRONIC ILLNESS

As recently reported, hospital-specific profiling is possible because most Medicare enrollees with serious chronic illnesses tend to use the same hospital throughout the course of those

illnesses.[11] For this study, the populations were therefore defined by assigning each patient to the hospital he or she most frequently used during the two years prior to death. For comparison, 77 institutions rated by *U.S. News & World Report* in 2001 as the nation's "best" hospitals for treating geriatric care, heart disease, cancer, and pulmonary disease were selected. Most of these hospitals are well-known academic medical centers.

These institutions' management styles were profiled using several measures that applied specifically to patients' last six months of life. These included: average number of days spent in the hospital during that time, average number of days spent in intensive care units (ICUs), average number of physician visits, percent of patients who see 10 or more physicians, percent of patients who die in ICUs, Medicare spending, and physician labor inputs.

Although selected for their reputations for high-quality care, these hospitals differed remarkably amongst themselves in the way they managed severely ill Medicare patients. This was often true even among hospitals in the same state or city.

Average Number of Days Spent in Hospitals
During the last six months of life, the number of days spent in hospitals ranged from 9.4 to 27.1 per decedent (Figure 11).[12] Patients assigned to the three academic medical centers in Manhattan were at the upper end—they had the highest patient day rates among the 77 hospital cohorts. Patients loyal to New York University (NYU) Medical Center spent almost a month in the hospital, while those assigned to Mount Sinai and New York–Presbyterian hospitals spent 22.8 and 21.6 days, respectively.[13] But among the four medical centers in California, there were striking differences in patterns of utilization. The average number of hospital days among patients assigned to the Cedars–Sinai Medical Center in Los Angeles was 21.3, very nearly the same as the New York teaching hospitals and more than twice the average for Stanford University Hospital, where decedents spent an average of 10.1 days of their last six months of life. Patients assigned to the University of California, Los Angeles (UCLA) Medical Center spent 16.1 days there, 24 percent fewer than patients at Cedars–Sinai—but 40 percent *more* days than among those at its sister organization, the University of California, San Francisco (UCSF) Medical Center (11.5 days).

Hospitals showing high rates of utilization among cohorts with one chronic condition tended to have high rates for cohorts with other chronic conditions. For example, the average number of days in the hospital for patient cohorts with congestive heart failure (CHF) and cancer were highly correlated (R^2 = .64) even though, on average, cancer patients tend to be hospitalized less often (Figure 12). There were similar

correlations between the rates of hospitalization for chronic obstructive pulmonary disease (COPD) and CHF, and between rates of hospitalizations for COPD and cancer. In other words, *the most important influence on the risk of spending time in the hospital was the hospital to which the patient was assigned, not whether they had cancer, CHF, or COPD.*

Also analyzed were racial differences in end-of-life care at the 50 "best" hospitals with 100 or more black patients. At the same hospital (controlling for case mix), black patients tended to use slightly more care than white patients—as evidenced by the predominance of dots above the 45-degree "equality" line in Figure 13. Hospital days among blacks—as among whites—varied by a factor of about 2.5 among the 50 hospitals, and the rates were highly correlated (R^2=.75). In other words, *what really mattered in determining the risk of hospitalization was not race but the hospital where most of the care was received.*

Why is so much of the variation in days in hospital explained by the hospital itself, rather than the illness that patients have or their relative need (as indicated by ethnicity)? Patients with CHF, COPD, and cancer are quite sick, particularly during the terminal phases of their illness, and physicians find it easier to manage these patients' often complex patterns of care in the hospital. Meanwhile, hospitals (and regions) with greater numbers of hospital beds per number of loyal patients have more opportunity to admit sick patients and to keep them in the hospital for longer periods. While blacks have slightly higher use rates than whites (perhaps reflecting blacks' relative lack of alternatives to hospital care), the effect on hospitalization rates of the particular hospital to which patients are loyal is much stronger than the effect of ethnicity.

Average Number of Days Spent in Intensive Care

During the last six months of life, the number of days spent in ICUs ranged from 1.6 to 9.5 days per decedent (Figure 14). The UCLA and Cedars-Sinai hospitals were near the top of the distribution, with 9.2 and 7.0 days, respectively. It is noteworthy that patients loyal to UCLA spent 3.5 times more days in intensive care than patients assigned to its sister hospital, UCSF (2.6 days). Stanford's use of ICU beds was 1.6 times greater than UCSF's. There were equally interesting contrasts in Manhattan. NYU Medical Center patients spent an average of 6.7 days in ICUs, 2.4 times more than patients loyal to Mount Sinai (2.8 days), while New York-Presbyterian patients, at 4.5 days, spent 1.6 times more days in the ICU than patients loyal to Mount Sinai Hospital.

As was the case with days in the hospital, days in ICUs were highly correlated among patients with different chronic illnesses and socioeconomic circumstances.[14] It is unclear, however, how hospitals such as Cedars-Sinai, UCLA, and NYU come to depend

so much on ICU beds in their care management plans while others, such as Mount Sinai and UCSF, get by with so much less.

Average Number of Physician Visits at End of Life
Physician visits during the last six months of life ranged from 17.6 to 76.2 per decedent among the 77 hospitals (Figure 15). As was the case with patient days, NYU Medical Center topped the list with an average of 76.2 visits. Patients loyal to Mount Sinai Hospital had an average of 53.9 visits, while New York-Presbyterian patients had 40.3. There were, again, striking differences among California teaching hospitals. Stanford (22.6 visits) and USCF (27.2) were at the lower end of the distribution. Cedars-Sinai was near the top, with 66.2 visits per decedent, almost three times greater than the average among patients loyal to Stanford. UCLA visits rates (43.9 per decedent) were 61 percent higher than UCSF rates, and 93 percent higher than rates among patients loyal to Stanford, but 34 percent lower than rates among patients loyal to Cedars-Sinai.

Patients who spend more days in hospitals receive more physician visits, as shown by the strong association (R^2 = .60) in Figure 16. The basis for this association is probably that referrals and revisits are much more easily scheduled when the patient is in the hospital. Similarly, on a given hospital day, patients are likely to be visited by several physicians, so the more days patients spend in hospitals, the more opportunities there are for visits.

Percentage of Patients at End of Life Seeing 10 or More Physicians
The proportion of patients who saw 10 or more physicians in their last six months of life varied from less than 17 percent to more than 58 percent (Figure 17). Mount Sinai and NYU were at or near the top of the distribution: 58.5 percent and 57.1 percent, respectively, of patients assigned to these hospitals saw 10 or more physicians. At New York-Presbyterian, the rate was 37.7 percent. Among the California hospitals, those located in Los Angeles were rather similar to those in New York: among patients loyal to UCLA and Cedars-Sinai, 50.9 percent and 48.2 percent, respectively, saw 10 or more physicians during their last six months of life. By contrast, among patients loyal to UCSF and Stanford, only 30.3 and 23.1 percent of patients, respectively, saw 10 or more physicians.

Patients who received most of their care from health care organizations that perform on the high end of this measure may suffer from lack of continuity of care—from what is sometimes called "ping-ponging" or "multiple-referral syndrome." Under such circumstances, lots of physicians get involved in care but no one is responsible for its

coordination. The inverse association between percentage of physicians involved in caring for chronically ill patients and scores on quality measures (e.g., percent in need who get effective care) is consistent with this interpretation.[15]

Percentage of Patients Who Die in ICUs

Another perspective on the quality of care is the quality of death, which ideally should be as free as possible from overly aggressive, futile care. However, there are striking differences among academic medical centers in the chance of dying in an ICU, which—for better or worse—has come to symbolize such an undesirable option.[16] About one-third of patients who were loyal to Cedars-Sinai, the UCLA Medical Center, and the NYU Medical Center died as hospital inpatients under treatment protocols that included at least one admission to an ICU (Figure 18). Only about 20 percent of patients loyal to UCSF, Stanford University Hospital, and Mount Sinai Hospital were so treated. These differences in care intensity need to be evaluated in light of Fisher's results, already discussed, which show that regions and academic medical centers with high rates of care do not have better health outcomes. *Greater intensity of terminal care, with its negative impact on the quality of dying, is thus not a price the dying must pay to ensure overall greater survival rates.*

Medicare Spending

The importance of dealing with unwarranted variation in the use of supply-sensitive care is underscored by our studies showing that this category of care "explains" most of the variations in per capita spending among regions. Per-enrollee Medicare spending varies almost threefold among hospital-referral regions and academic medical centers, with greater spending being the result in large measure of local providers having higher utilization rates for supply-sensitive care: more physician visits, hospitalizations, stays in ICUs, and diagnostic testing and imaging. Regions and academic medical centers with greater overall spending rates do not, however, have higher quality of care. *In view of the Fisher findings, the problem is not underuse in low-rate regions and hospitals; it is overuse and inefficiency in high-rate regions.*

It is important to note that the patterns of practice and Medicare spending in the last six months of life are an indicator of the relative intensity of care delivered to the chronically ill during previous stages in the progression of their disease. This is evident from the high correlations between Medicare spending during those last six months and spending for the same patient cohort during earlier periods. For example, the overall average per-decedent spending for Part A inpatient care and Part B physician and laboratory services for the 77 *U.S. News & World Report* "best" hospitals in the last six months of life was $22,000, more than five times higher than the $3,900 average for the

same cohorts in the 18th to 24th months prior to death. However, Medicare program spending varied almost threefold among the 77 hospitals cohorts, from $11,500 to $37,200 per decedent during the last six months and from $2,200 to $8,100 during the 18th to 24th months prior to death. The spending patterns were very highly correlated (R^2 = .79) (Figure 19).

Spending levels for care in the last six months of life provide a case-mix-adjusted profile of the efficiency of a health care organization in managing chronic illness—one that is untainted by differences in illness severity.

Physician Labor Inputs

Figure 20 examines the amount of physician labor, measured in terms of standardized full-time equivalents (FTEs) invested in the care of 67 hospital cohorts.[17] The data reveal large variations in the way physician labor is used in treating chronic illnesses. During the last six months of life, labor input of medical specialists range from 1.8 to 15.5 FTEs per 1,000 decedents, while inputs of primary care range from 2.4 to 10.4 FTEs per 1,000. Among the California and New York cohorts, the combined input rates for primary physicians and medical specialists during the last six months of life ranged from 8.4 FTEs for Stanford to 24.6 FTEs for NYU, a threefold range in variation. The combined inputs to Cedars-Sinai and Mount Sinai cohorts were 20.7 and 16.4 FTEs per 1,000 decedents, respectively. UCLA used 59 percent more physician labor than UCSF. Note also from the data of Figure 20 the wide range of variation in reliance on primary care physicians versus medical specialists. For example, the ratio of medical specialist to primary care input rates for USCF was 0.67 while for UCLA it was 2.84.

Measures of resource inputs such as these are important for the population-based management of care. Heretofore they have been available only to clinicians and managers of HMOs such as Kaiser Permanente. By making them available for fee-for-service organizations, the hope is to stimulate accountability for capacity—an essential component of any strategy to reduce the overuse of supply-sensitive care.[18]

ACHIEVING SUSTAINABLE IMPROVEMENT IN QUALITY AND EFFICIENCY

The Dartmouth Atlas Project recently made hospital-specific information available for California, and plans to do subsequent releases regarding other parts of the United States. The simple availability of information on the relative efficiency of specific health care organizations in managing chronic illness—what Arnold Milstein, medical director of the Pacific Business Group on Health, has called "longitudinal efficiency"—could prove

beneficial. It may stimulate payers to reexamine their provider networks (which traditionally have been based on unit price, not volume times price) and motivate employers to steer their employees toward efficient hospitals.

Assuming that the trends seen for Medicare apply also to other payers, successful redesign along these lines would lead to net savings for employers and payers who can flexibly direct their patients to such providers.[19] It would also ensure the profitability of those health plans participating in Medicare Advantage (Medicare managed care); they could make deals to send their patients to physician groups using hospitals with spending levels below the regional average.

Ironically, unless it too could join in directing patients to efficient providers, traditional Medicare stands to lose. If commercial payers steered patients away from the high-cost providers, the population loyal to such providers would shrink but available resources would not. This would result in yet higher utilization rates and costs for supply-sensitive care, possibly worsening outcomes among the chronically ill Medicare patients who remained loyal to such providers.

The availability of provider-specific estimates for the actuarial costs of care discussed above is a step in that direction. They may provide opportunity for new thinking in the design of "budget-neutral" reimbursement strategies, such as "partial capitation," that would provide preferred providers with budgets to help compensate them for losses in revenue associated with reductions in inpatient care.

Also recommended is a demonstration project between the Centers for Medicare and Medicaid Services and progressive health care organizations that share the goal of reducing unwarranted variation in all three categories of care.[20] As experience is gained and the quality of care improves, additional incentives might be put in place—such as rewarding managers who use benchmarks from efficient providers in the recruiting of medical personnel and the construction of facilities—to further enhance population-based management. The measures of workforce labor input, reviewed above, could be useful for this purpose.

In the long run, the most challenging problem will be finding mechanisms to clear regional markets of excess capacity. While special deals made with forward-thinking providers may well result in models of how to deliver care that is simultaneously of high quality and low cost, strategies to ensure that all Medicare patients are served by such hospitals remain elusive. If Medicare administrators were willing and able, however, to
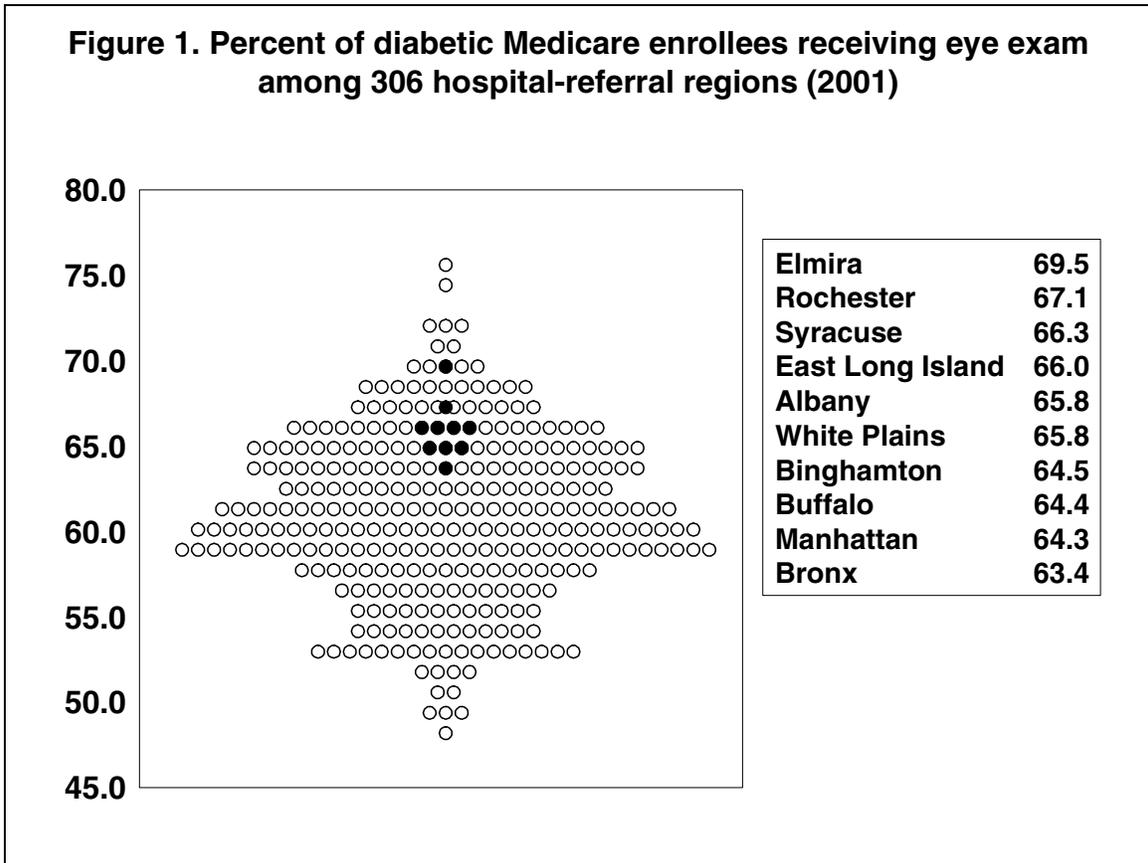
take steps to select providers on the basis of quality and efficiency—and other payers were willing to play by similar rules—this would serve as a wakeup call to the provider community. Presumably, it would result in accelerated change throughout the nation's health care markets.

SUMMARY: THE PROBLEM OF UNWARRANTED VARIATION
The economic and clinical implications of practice variation, and the opportunities and strategies for reform, depend on the category of care. Having reviewed examples of effective care, preference-sensitive care, and supply-sensitive care, and having discussed the causes of unwarranted variation, this report finds that:

- Most kinds of effective care—beta-blockers for heart attack patients, for example, or screening of diabetics for early signs of retinal disease—are characterized by underuse. Its causes include discontinuity of care (worsened when more physicians are involved in the care) and lack of infrastructure to ensure outreach and the timely use of these services. Pay-for-performance strategies could reduce such underuse.

- Preference-sensitive care, in which treatment options involve significant tradeoffs that should be based on the patient's own values, tends to be misused. The causes of this misuse include failure to accurately communicate the risks and benefits of the alternative treatments and the failure to base choice of treatment on the patient's opinion rather than those of others. Adjustment of economic incentives to reward adopters of shared decision-making could lead to a reduction in such unwarranted variation.

- Supply-sensitive care is overused, particularly in the management of chronic illness. The causes include overdependence on acute hospital care and lack of infrastructure to support continuous management of chronically ill patients in other care settings. Hospital-specific measures that profile performance in managing chronic illness can help identify efficient providers. Moreover, pay-for-performance strategies, along with related strategies to reward efficient providers and pay for chronic illness management infrastructure, could promote reform.

Figure 1. Percent of diabetic Medicare enrollees receiving eye exam among 306 hospital-referral regions (2001)

Elmira	69.5
Rochester	67.1
Syracuse	66.3
East Long Island	66.0
Albany	65.8
White Plains	65.8
Binghamton	64.5
Buffalo	64.4
Manhattan	64.3
Bronx	63.4

Each dot represents the score of a distinct New York State region on a quality measure for diabetic care. This score is the percentage of diabetics who received the medically necessary care—an annual eye examination—in the region relative to the total number of diabetics living there. The figure highlights in solid black the location of hospital-referral regions within New York City.

Figure 2. Rates of four orthopedic procedures among Medicare enrollees in 306 hospital-referral regions (2000–01)
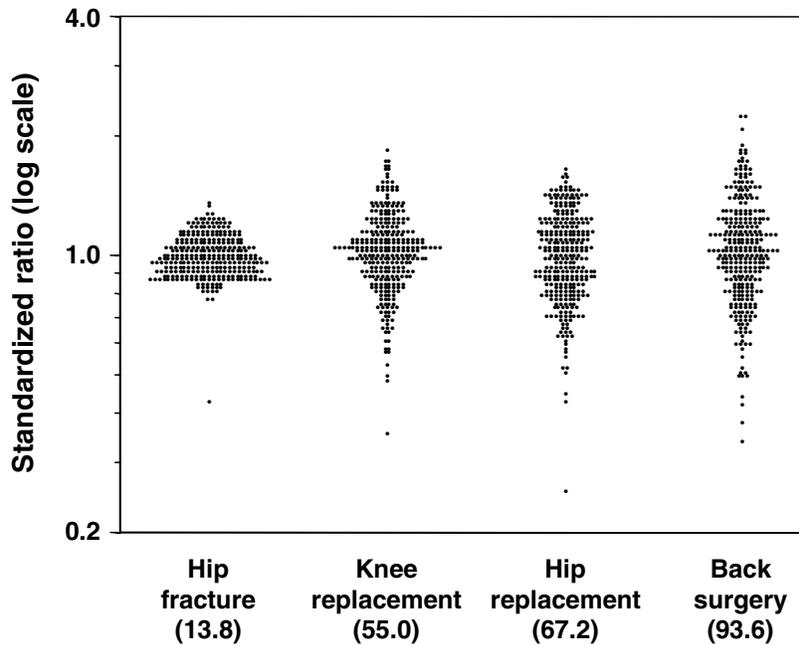
This figure profiles the pattern of variation among 306 hospital-referral regions regarding four orthopedic procedures: hip fracture repair, knee replacement, hip replacement, and back surgery. Each dot represents one of the 306 regions. The rates—whereby the numerator is the region's number of patients with the indicated procedure and the denominator is number of enrollees in traditional Medicare living in the region—are expressed as the ratio to the U.S. average (plotted on a log scale). The numbers in parentheses are the systematic components of variation, measures that allow comparisons of variation among procedures with different mean rates. In the regions represented here, knee replacement is about four times more variable than hip fracture repair, and back surgery is almost seven times more variable.

20

Figure 3. Surgical signatures of four Florida hospital-referral regions compared to the Manhattan hospital-referral region (2000–01)

Ratio to Manhattan HRR

	Fort Myers	Sarasota	Fort Lauderdale	Miami
Knee replacement	3.04	2.48	1.80	1.26
Hip replacement	2.04	1.92	1.72	0.89
Back surgery	3.37	3.24	2.06	1.39

■ Knee replacement　　■ Hip replacement　　□ Back surgery
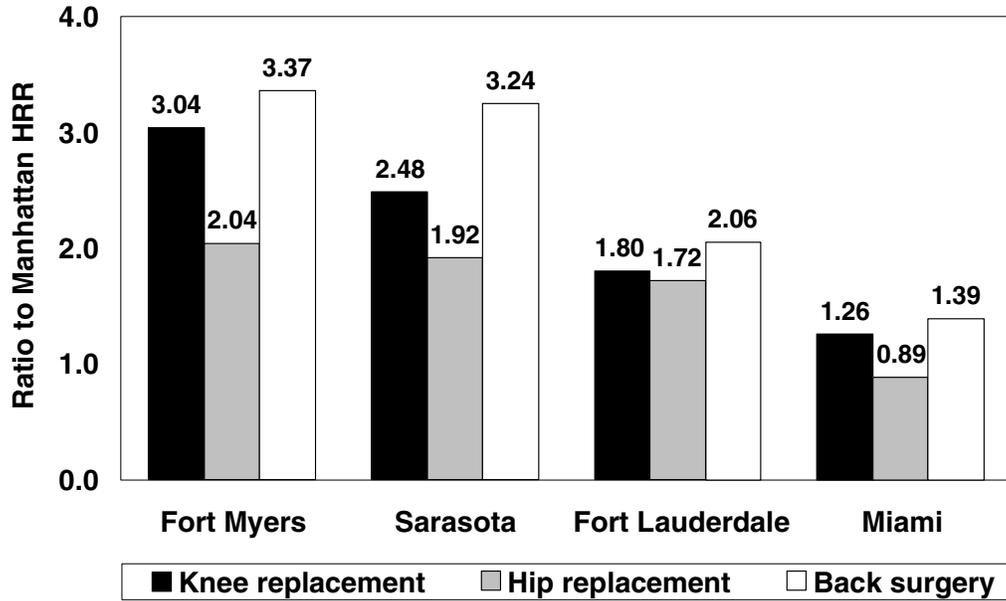
This figure profiles the rates of knee replacement, hip replacement, and back surgery in four South Florida medical communities. Each rate is expressed as a multiple of the corresponding Manhattan rate. For example, the rate of knee replacement in Fort Myers is 3.04 times greater than that of Manhattan. All rates are age-, sex-, and race-adjusted.

Figure 4. Rates of orthopedic procedures among three Florida hospital-referral regions relative to the Manhattan hospital-referral region

	92–93	94–95	96–97	98–99	00–01	Number of cases above or below Manhattan HRR
Fort Myers (165,025)						
Back surgery	4.31	4.07	3.51	3.30	3.37	7,246
Knee replacement	3.43	3.02	2.43	2.57	3.04	7,099
Hip replacement	2.19	2.03	1.81	2.00	2.04	2,689
Fort Lauderdale (284,081)						
Back surgery	2.26	2.18	2.22	1.85	2.06	5,680
Knee replacement	1.75	1.80	1.61	1.59	1.80	5,124
Hip replacement	1.81	1.61	1.74	1.70	1.72	3,618
Miami (174,781)						
Back surgery	1.16	1.12	1.26	1.30	1.39	868
Knee replacement	1.38	1.39	1.29	1.16	1.26	1,422
Hip replacement	1.09	1.01	1.00	0.93	0.89	–56

(2001 Medicare population)

This table profiles the rates, during two-year periods over the decade from 1992 to 2001, for knee replacement, hip replacement, and back surgery in three South Florida medical communities. Rates are expressed as ratios of local to Manhattan rates during the corresponding period, and these ratios are quite consistent from year to year. Accumulating over the decade, the "excess" number of cases (compared to what would have obtained had Manhattan rates prevailed locally) for all three procedures reached 17,000 operations on the Medicare residents of Fort Myers and 14,400 for residents of Fort Lauderdale.

Figure 5. Use of physician services and hospitalizations for chronic conditions among Medicare enrollees in 306 hospital-referral regions (1995–96)
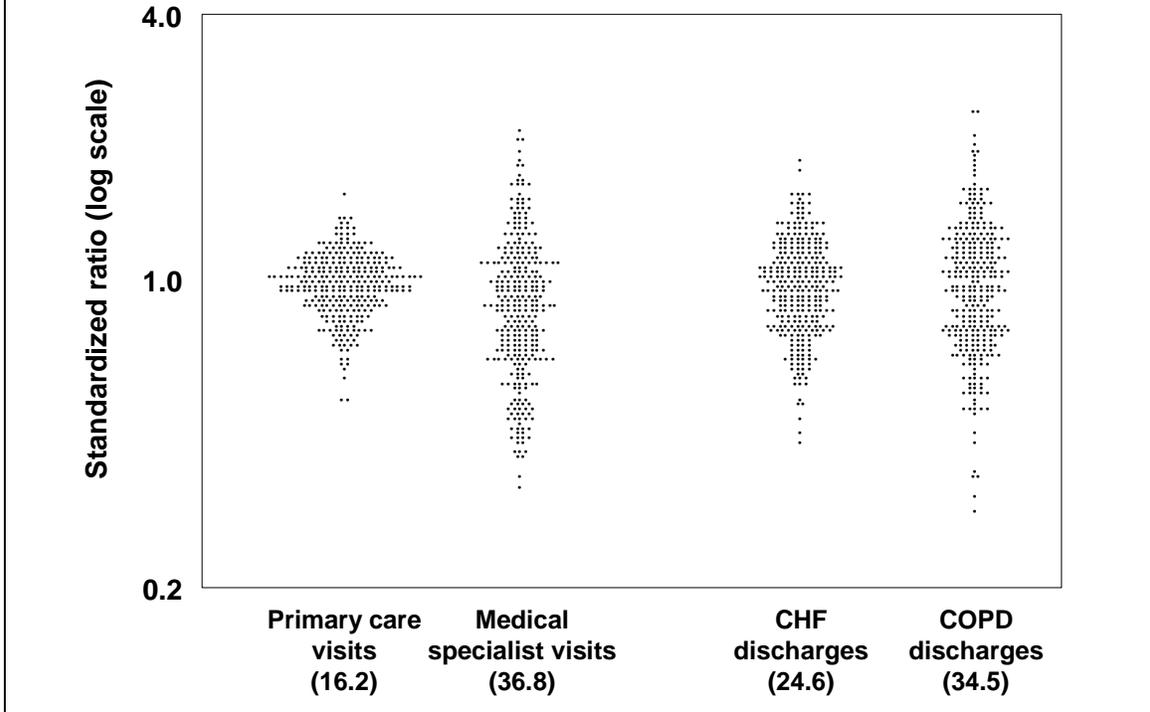
This figure profiles the pattern of variation (age-, sex-, and race-adjusted) for selected supply-sensitive services. Each dot represents one of the 306 regions. The numbers in parentheses are each service's coefficient of variation. Primary care visits vary about threefold and demonstrate the least variation. Visits to medical specialists vary more than fivefold, as do discharges for chronic obstructive pulmonary disease (COPD); and congestive heart failure (CHF) discharges vary about fourfold.

Figure 6. Patient days in hospital during the last six months of life among Medicare decedents in 306 hospital-referral regions (2001)

Manhattan	20.0
East Long Island	19.0
Bronx	18.5
White Plains	16.3
Elmira	13.9
Syracuse	12.6
Albany	12.1
Buffalo	11.8
Binghamton	11.8
Rochester	10.6
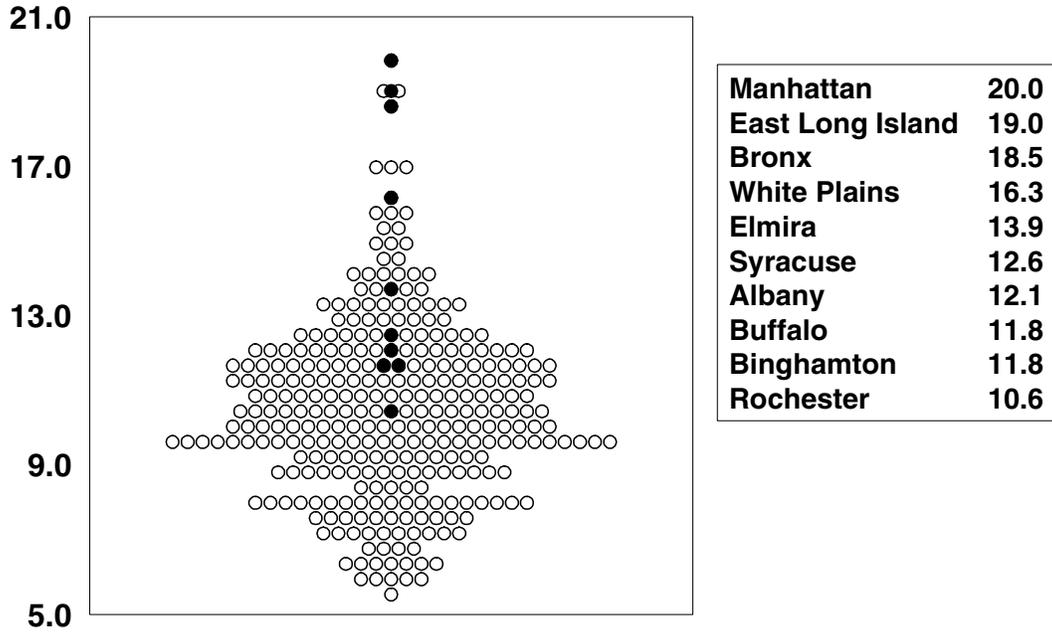
The use of hospitals for medical conditions is particularly intense during the last few months of life, but there is striking variation among regions. This figure gives the distribution in rates among the 306 regions for days spent in the hospital by resident Medicare enrollees during the last six months of their life.

Figure 7. Association between hospital beds per 1,000 (1996) and discharges per 1,000 (1995–96) among Medicare enrollees in 306 hospital-referral regions
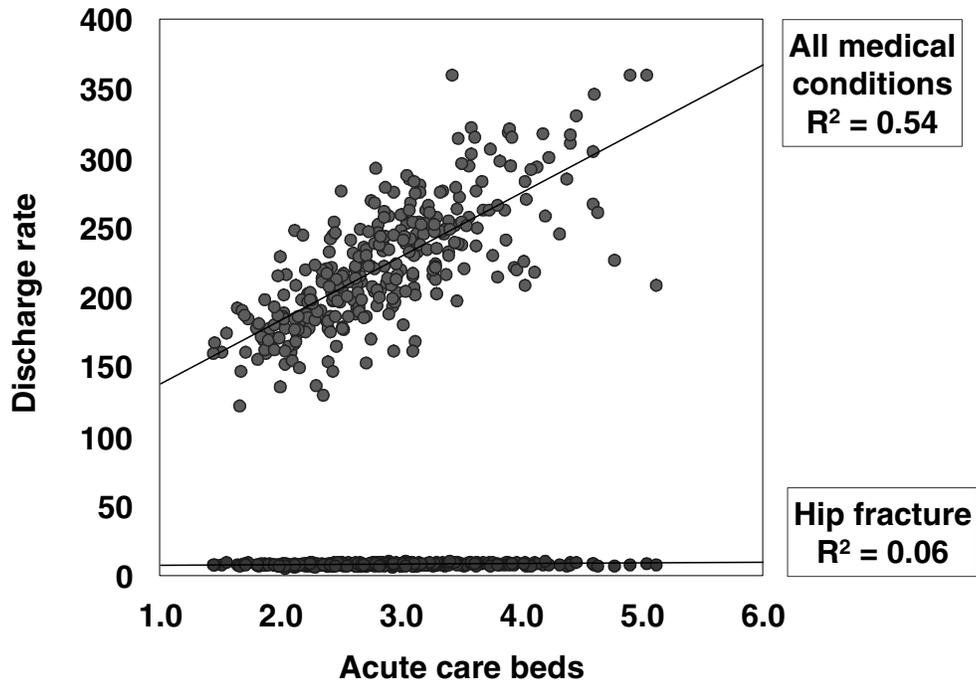
This figure shows the association between supply of hospital beds and the hospitalization rate for medical (nonsurgical) conditions. More than half of the variation in discharge rates is associated with bed capacity. By contrast, hospitalization for hip fracture—one of the few conditions for which the pattern of variation is determined by the incidence of illness—shows little correlation with resource supply.

**Figure 8. Association between cardiologists and
visits per person to cardiologists among Medicare enrollees (1996):
306 hospital-referral regions**

$R^2 = 0.49$

This figure illustrates the relationship between the number of cardiologists per 100,000 and the number of visits per person to cardiologists among the 306 regions. About half of the variation is associated with supply.

Figure 9. Total physician visits during the last six months of life among Medicare decedents in 306 hospital-referral regions (2001)

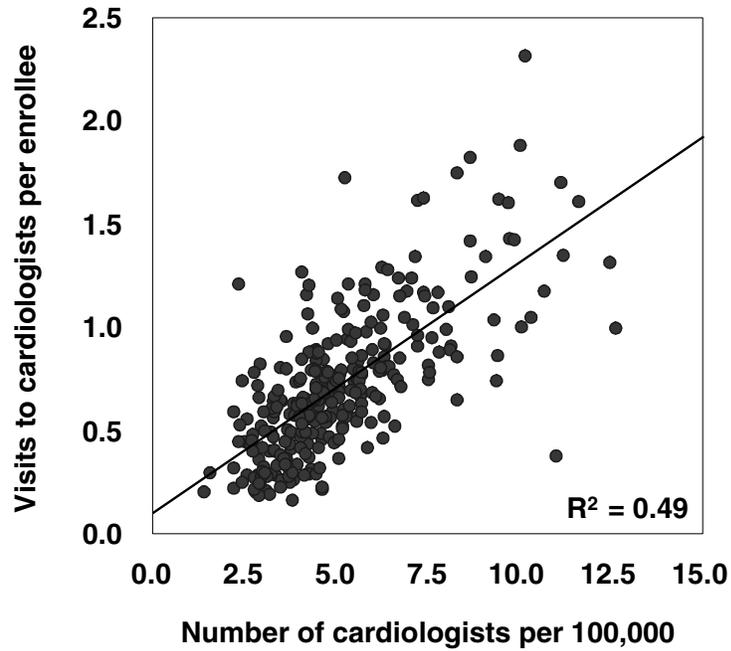|	
East Long Island	**39.4**
Manhattan	**37.2**
White Plains	**35.3**
Bronx	**32.9**
Elmira	**24.5**
Albany	**24.5**
Buffalo	**20.7**
Syracuse	**20.2**
Binghamton	**19.8**
Rochester	**16.1**

Physician visits are particularly frequent during the last few months of life, but there is striking variation among regions. This figure gives the distribution in rates among the 306 regions.
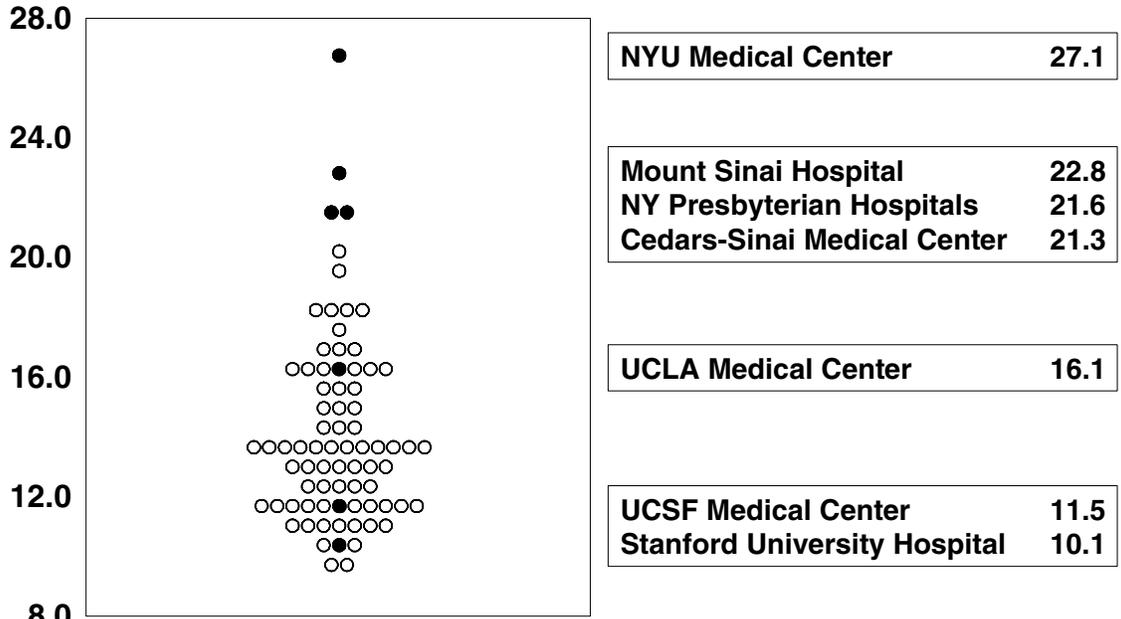
**Figure 10. Per-capita resource inputs and health outcomes:
Ratio between high and low quintiles in spending among
306 hospital-referral regions**

Resource inputs		Cohort health outcomes		
Medicare spending	1.61	Death	R.R.	95% CL
Hospital beds (1000)	1.32			
		Hip fracture	1.019	1.001–1.039
Physician supply*		Colon cancer	1.012	1.018–1.094
All physicians	1.31	Heart attack	1.052	1.018–1.094
Medical specialists	1.65			
General internists	1.75	Functional status: No difference		
Family practice	0.74	Satisfaction: No difference		
Surgeons	1.37	Access: Worse		

* per 10,000

This table, adapted from the Fisher study, compares the regions in the highest quintile of Medicare spending with those in the lowest quintile. Results on the left indicate the level of resource inputs and on those on the right document the health care outcomes for local patients. See text for explanation.

Figure 11. Days spent in hospitals during the last six months of life among patients who received most of their care at one of the 77 "best" U.S. hospitals

NYU Medical Center	27.1

Mount Sinai Hospital	22.8
NY Presbyterian Hospitals	21.6
Cedars-Sinai Medical Center	21.3

UCLA Medical Center	16.1

UCSF Medical Center	11.5
Stanford University Hospital	10.1

Each dot represents the average number of days per person spent at one of the 77 hospitals during the last six months of life. The figure highlights in solid black the academic medical centers located in Manhattan and in California.

**Figure 12. Association between hospital days
for cancer and for CHF patients during the last six months of life
among the 77 "best" U.S. hospitals**
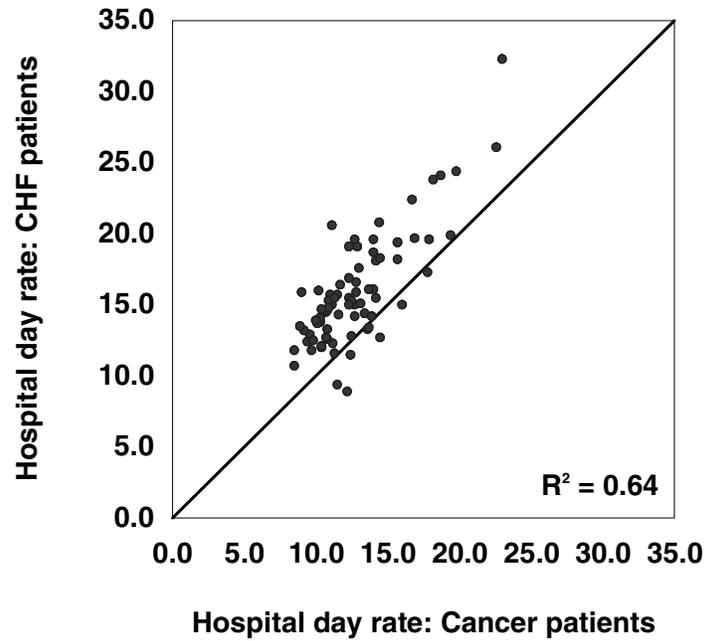
This figure examines the relationship between average number of days in the hospital for patient cohorts with congestive heart failure (CHF) and those with solid-tissue cancers. Each dot represents the rates for patients assigned to a given hospital.

Figure 13. Association between hospital days for black and non-black patients during the last six months of life among the 50 "best" hospitals

This figure examines utilization rates among black (vertical axis) and non-black (horizontal axis) members of the patient cohorts for those "best" hospitals with 100 or more black patients.

Figure 14. Days spent in intensive care during the last six months of life among patients receiving most of their care at one of the 77 "best" U.S. hospitals
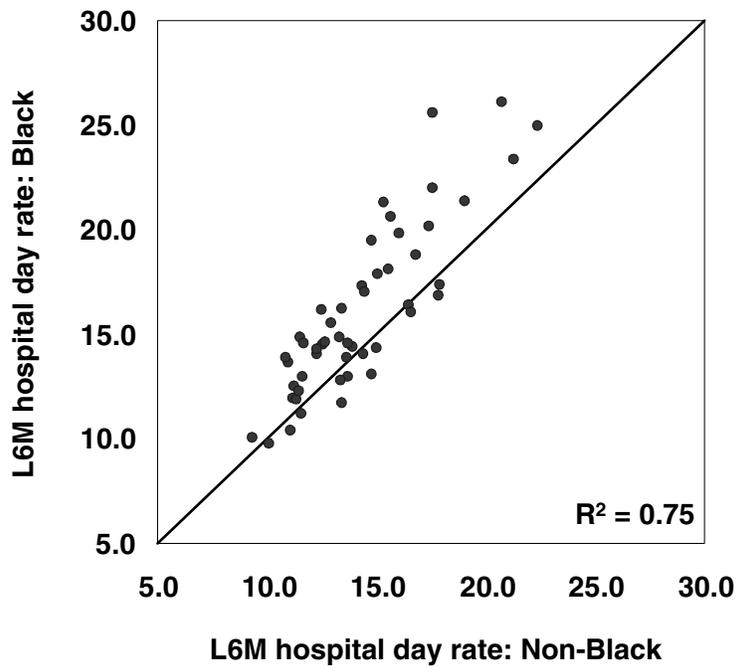
UCLA Medical Center	9.2
Cedars-Sinai Medical Center	7.0
NYU Medical Center	6.7
NY Presbyterian Hospital	4.5
Stanford University Hospital	4.3
Mount Sinai Hospital	2.8
UCSF Medical Center	2.6
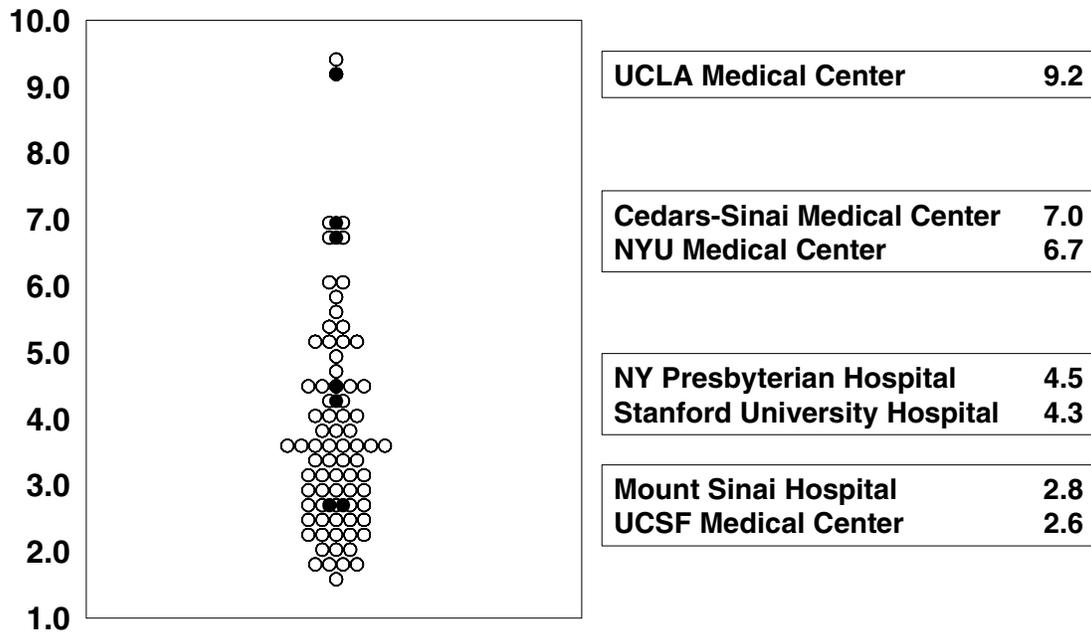
Each dot, representing one of the 77 hospital cohorts, shows the average number of days spent in intensive care per person during the last six months of life.

Figure 15. Average number of physician visits per patient during the last six months of life who received most of their care at one of the 77 "best" U.S. hospitals
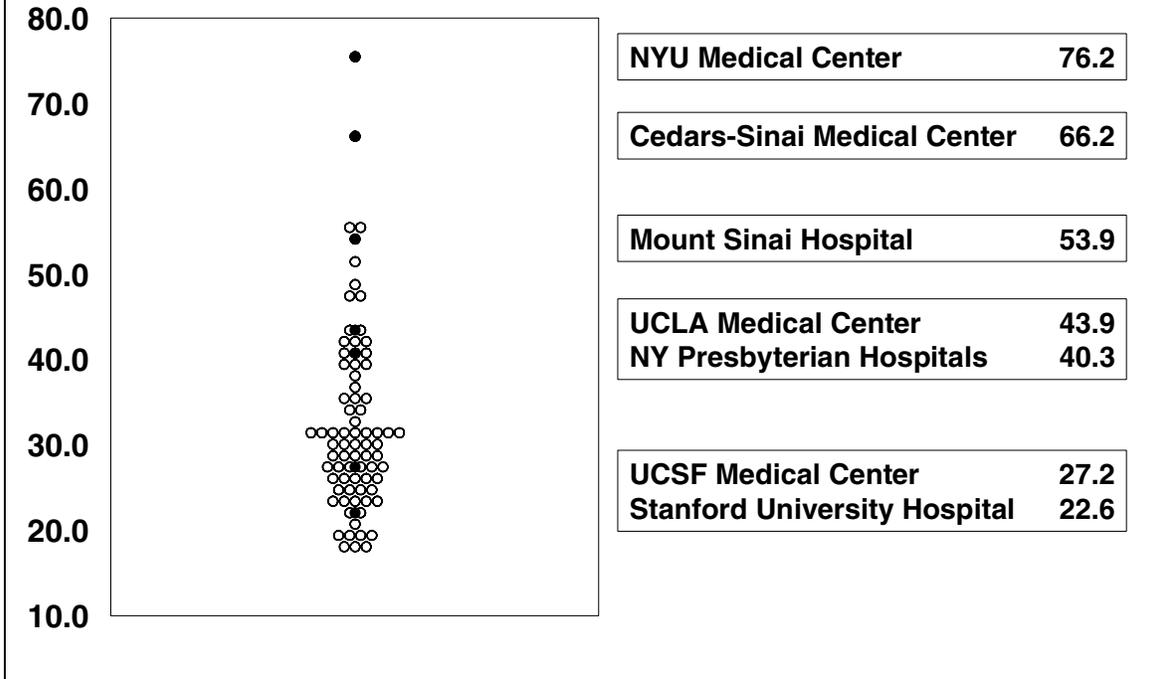
NYU Medical Center	76.2
Cedars-Sinai Medical Center	66.2
Mount Sinai Hospital	53.9
UCLA Medical Center	43.9
NY Presbyterian Hospitals	40.3
UCSF Medical Center	27.2
Stanford University Hospital	22.6

Each dot, representing one of the 77 hospital cohorts, shows the average number of physician visits per person during the last six months of life.

Figure 16. Association between hospital days and physician visits during the last six months of life among patients receiving most of their care at one of the 77 "best" U.S. hospitals

$R^2 = 0.60$

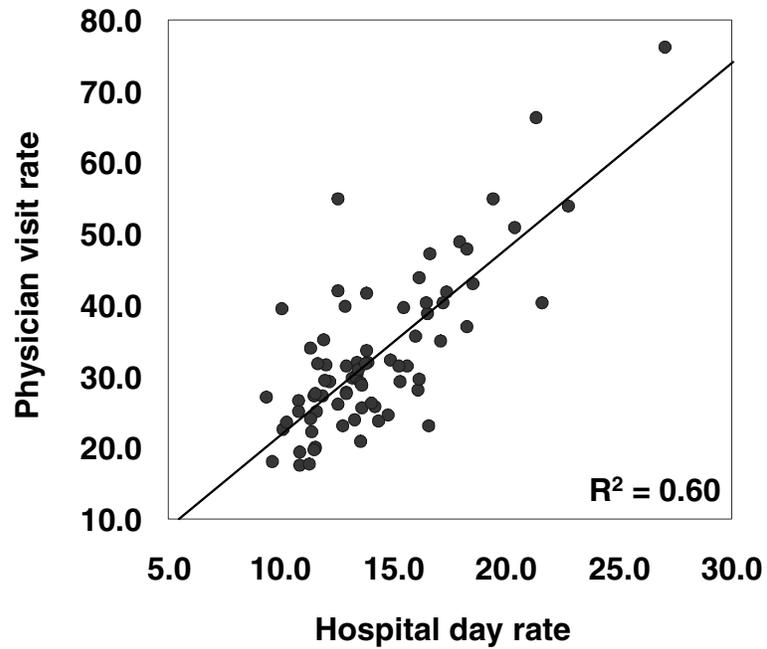This figure shows the association, during the last six months of life among patients of the 77 "best" U.S. hospitals, between days spent in the hospital per person and physician visits per person.

Figure 17. Percent of patients, receiving most of their care at one of the 77 "best" U.S. hospitals, seeing 10 or more physicians during the last six months of life

Mount Sinai Hospital	58.5
NYU Medical Center	57.1

UCLA Medical Center	50.9
Cedars-Sinai Medical Center	48.2

NY Presbyterian Hospital	37.7

UCSF Medical Center	30.3

Stanford University Hospital	23.1
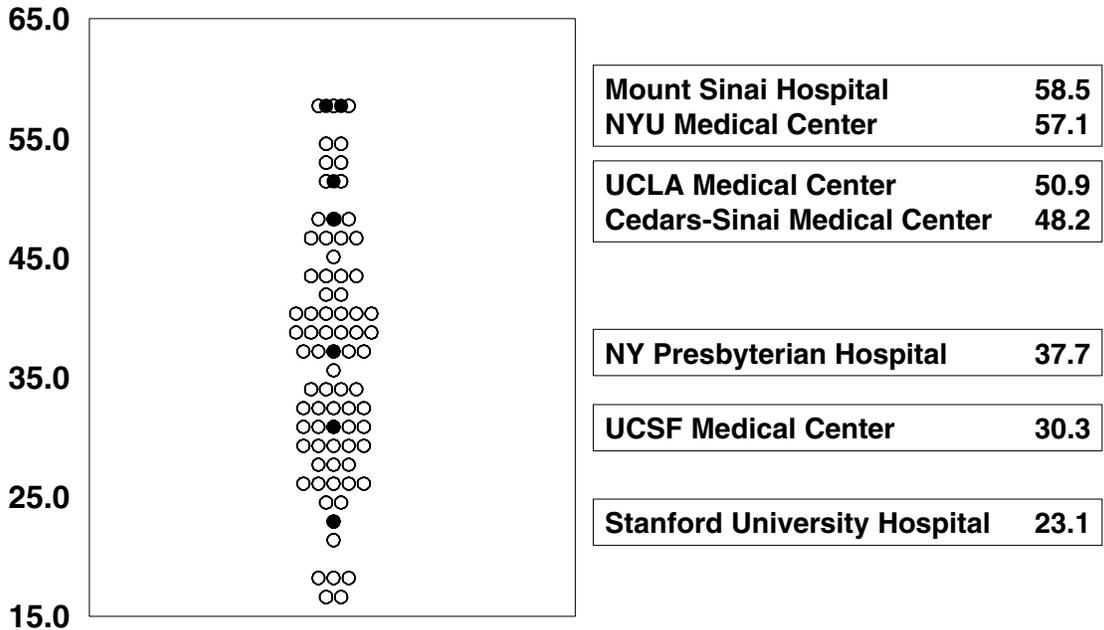
Each dot, representing one of the 77 hospital cohorts, shows the percent of patients who saw 10 or more physicians during the last six months of life. For example, 58.5 percent of the patients who were assigned to Mount Sinai Hospital saw 10 or more physician visits, while only 23.1 percent assigned to Stanford University Hospital did.

Figure 18. Percent of deaths associated with admission to intensive care units among patients receiving most of their care at one of the 77 "best" U.S. hospitals

Cedars-Sinai Medical Center	**36.8**

UCLA Medical Center	**32.9**
NYU Medical Center	**32.8**
NY Presbyterian Hospital	**28.5**

UCSF Medical Center	**22.3**
Stanford University Hospital	**21.3**
Mount Sinai Hospital	**20.0**
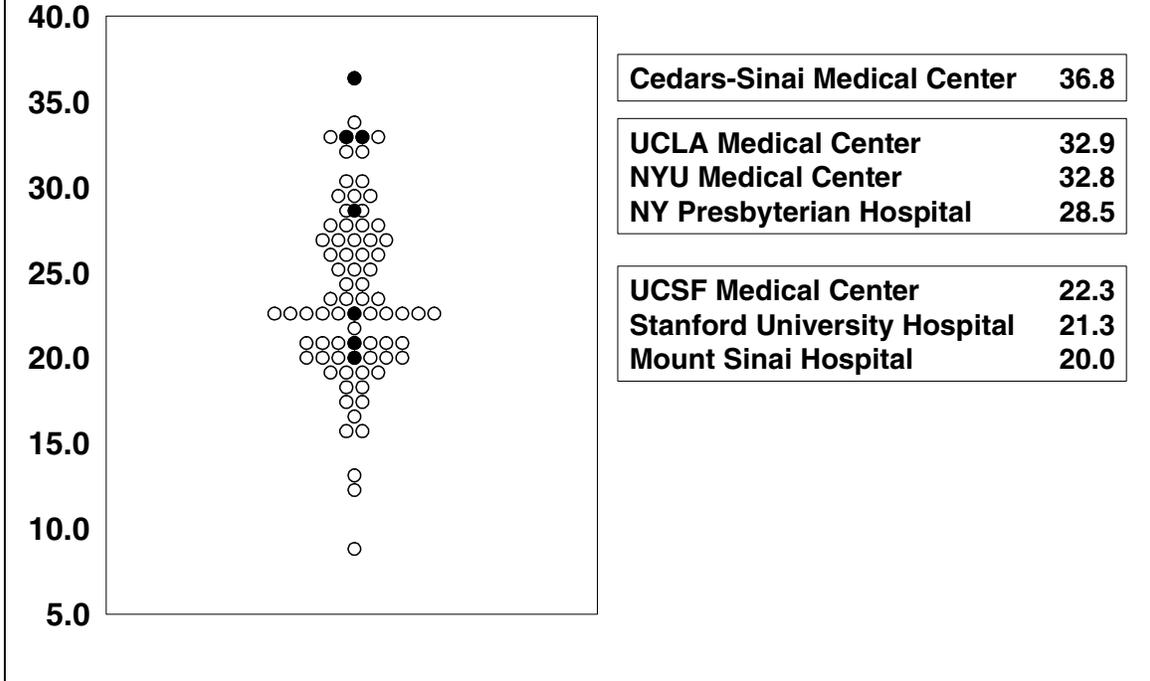
Each dot, representing one of the 77 hospitals, shows the percent of deaths associated with hospitalization in an intensive care unit.

Figure 19. Association between total Medicare payments 18–24 months and 0–6 months before death: 77 hospital cohorts (1999–01)

$R^2 = 0.79$

Total payments 0–6 months before death

Total payments 18–24 months before death
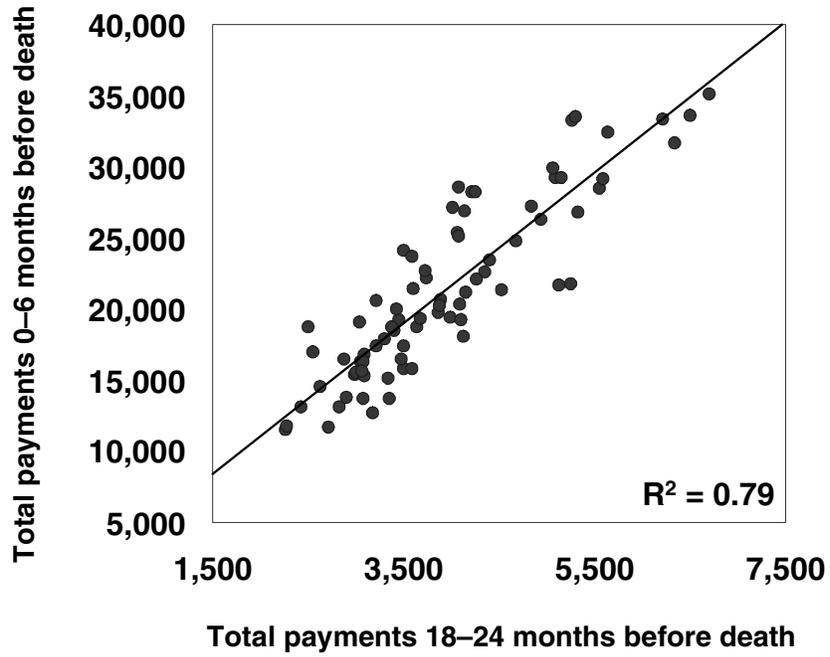
This figure correlates Medicare spending (Part A and Part B) per decedent during the last 6 months of life and during the period 18–24 months prior to death.

Figure 20a. Primary care S-FTE inputs per 1,000 Medicare decedents during the last six months of life among patients receiving most of their care at one of the 67 "best" hospitals

NYU Medical Center	9.1
Mount Sinai Hospital	7.8
New York Presbyterian	6.7
Cedars-Sinai Medical Center	6.5
UCSF Medical Center	5.5
UCLA Medical Center	3.8
Stanford University Hospital	3.8
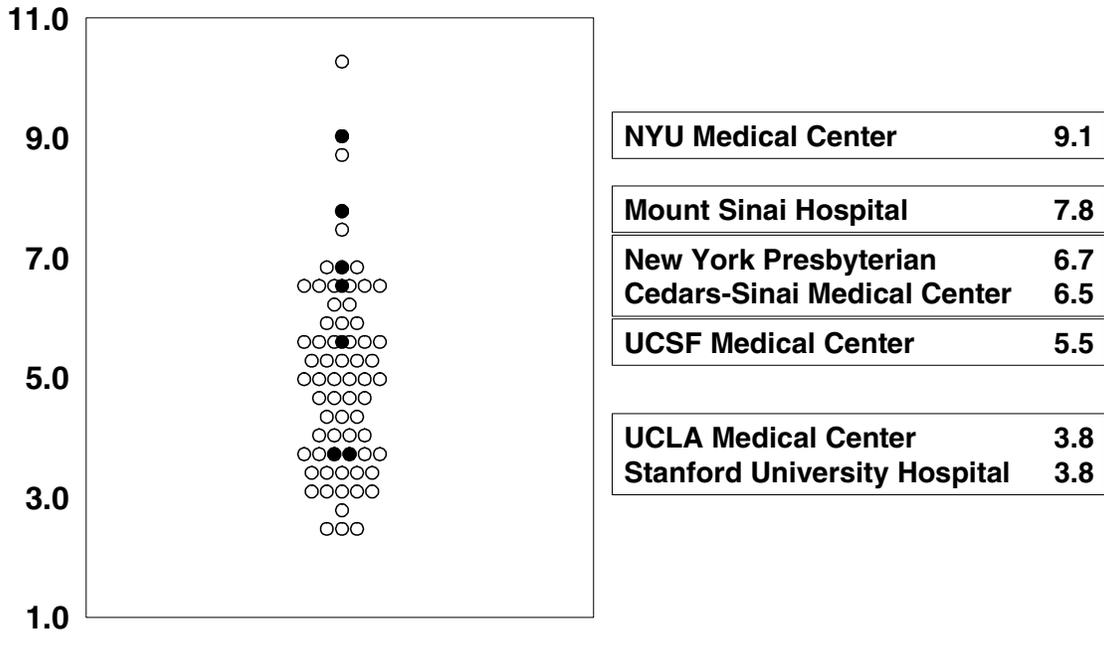
Figure 20b. Medical specialist S-FTE inputs per 1,000 Medicare decedents during the last six months of life among patients receiving most of their care at one of the 67 "best" hospitals

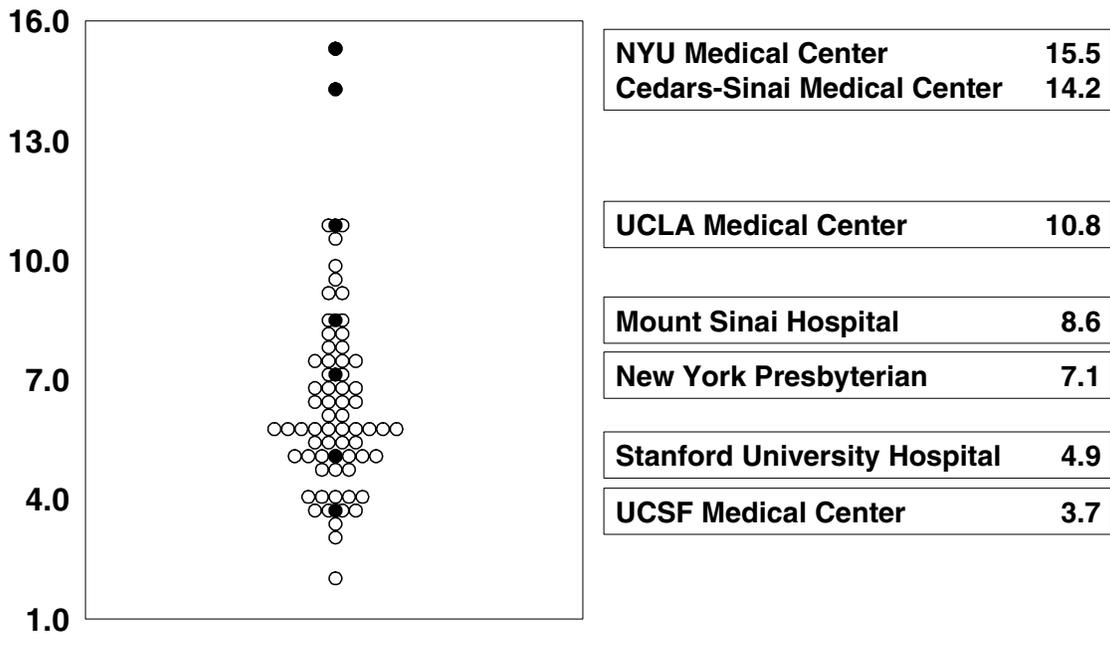|	
NYU Medical Center	15.5
Cedars-Sinai Medical Center	14.2
UCLA Medical Center	10.8
Mount Sinai Hospital	8.6
New York Presbyterian	7.1
Stanford University Hospital	4.9
UCSF Medical Center	3.7

The figure provides estimates of standardized full-time equivalent labor inputs for primary care physicians (20a) and medical specialists (20b) among 67 "best" hospital cohorts. See text for explanation.

NOTES

[1] E. A. McGlynn, S. M. Asch, J. Adams et al., "The Quality of Health Care Delivered to Adults in the United States," *New England Journal of Medicine* 348 (June 26, 2003): 2635–45.

[2] K. Baicker and A. Chandra, "Medicare Spending, the Physician Workforce, and Beneficiaries' Quality of Care," *Health Affairs* Web Exclusive (April 7, 2004): W4-184–W4-197, http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.184v1/DC1.

[3] A. M. O'Connor, H. A. Llewellyn-Thomas, and A. B. Flood, "Modifying Unwarranted Variations in Health Care: Shared Decision Making Using Patient Decision Aids," *Health Affairs* Web Exclusive (October 7, 2004): VAR–63–VAR 72, http://content.healthaffairs.org/cgi/content/full/hlthaff.var.63/DC2.

[4] G. A. Hawker, J. G. Wright, P. C. Coyte et al., "Determining the Need for Hip and Knee Arthroplasty: The Role of Clinical Severity and Patients' Preferences," *Medical Care* 39 (March 2001): 206–16.

[5] The absence of a strong association between the per capita supply of orthopedic surgeons and rates of knee replacement, hip replacement, and back surgery also applies to other surgical specialists with regard to procedures performed on the Medicare population. Although the supply of cardiovascular surgeons, cardiologists, urologists, general surgeons, and vascular surgeons vary by factors of more than three among regions, there is little association between the per-capita supply of those specialists and the rates of common procedures they perform. The R^2 statistic ranged from .00 (for the association between urologists per capita and transurethral prostatectomy) to .09 (relating the number of vascular surgeons to cases of lower-extremity bypass grafting).

[6] E. H. Wagner, P. Barrett, M. J. Barry et al., "The Effect of a Shared Decisionmaking Program on Rates of Surgery for Benign Prostatic Hyperplasia: Pilot Results," *Medical Care* 33 (August 1995): 765–70.

[7] In the author's experience, the impact of beds per capita on clinical decision making is subliminal. Clinicians are unaware of the threshold effect supply exerts on their decision making. This impression derives from interviews with clinicians practicing in Boston and New Haven, who were not aware of the 60 percent differences in hospital beds and hospitalization rates for medical conditions between their regions. Moreover, clinicians who had practiced in both communities were unaware that hospitalization rates were substantially different in the two settings in which they had practiced medicine.

[8] *British Medical Journal,* Annual *Clinical Evidence Concise 12* (December 2004). BMJ Publishing Group Ltd.

[9] "Roemer's Law" was named in honor of Milton Roemer (see reference from The Dartmouth Atlas of Health Care). In the early 1960s Roemer, a health services researcher interested in the use of hospitals, suggested that hospital beds, once built, will be used—no matter how many there are. The relationship between the capacity of the acute hospital sector (measured in beds per thousand residents of the local hospital-referral region) and the costs of care is an important illustration of that "law."

[10] E. S. Fisher, D. E. Wennberg, T. A. Stukel et al., "The Implications of Regional Variations in Medicare Spending: Part 1. The Content, Quality, and Accessibility of Care," *Annals of Internal Medicine* 138 (February 18, 2003): 273–87, http://www.annals.org/cgi/content/full/138/4/273; E. S. Fisher, D. E. Wennberg, T. A. Stukel et al., "The Implications of Regional Variations in Medicare Spending: Part 2. Health Outcomes and Satisfaction with Care," *Annals of Internal Medicine* 138 (February 18, 2003): 288–98, http://www.annals.org/cgi/content/full/138/4/288.

[11] D. E. Wennberg, E. S. Fisher, T. A. Stukel et al., "Use of Hospitals, Physician Visits, and Hospice Care During Last Six Months of Life Among Cohorts Loyal to Highly Respected Hospitals in the United States," *British Medical Journal* 328 (March 13, 2004): 607–11, http://bmj.bmjjournals.com/cgi/content/full/bmj;328/7440/607; D. E. Wennberg, E. S. Fisher, T. A. Stukel et al., "Use of Medicare Claims Data to Monitor Provider-Specific Performance Among Patients with Severe Chronic Illness," *Health Affairs* Web Exclusive (October 7, 2004): VAR-5–VAR-18, http://content.healthaffairs.org/cgi/content/abstract/hlthaff.var.5v1.

[12] Rates are based on all hospitalizations during the last six months of life, mostly spent in the hospitals to which the patients were assigned. Severe chronic illness was defined as complicated illness in 12 categories proposed by Iezonni and her colleagues (L. I. Iezonni, T. Heeren, S. M. Foley et al., "Chronic Conditions and Risk of In-Hospital Death," *Health Services Research* 29 (October 1994): 435–60). Rates are adjusted for age, sex, race, and type of chronic illness. For a full listing of results (including confidence limits), see http://www.dartmouthatlas.org.

[13] Because of the way the hospitals were coded, the experiences of New York Hospital and Presbyterian Hospital could not be examined separately. The estimate is the weighted average for the two organizations.

[14] See J. Skinner, A. Chandra, D. Staiger, J. Lee, and M. McClellan, "Mortality After Acute Myocardial Infarction in Hospitals That Disproportionately Treat Black Patients," *Circulation*, October 2005 112(17):2634–2641.

[15] Baicker, "Medicare Spending," 2004.

[16] The measure is an approximation in that exact date of discharge from the ICU and date of death were not matched.

[17] These measures were available for only 67 of the 77 hospitals. For information on how these measures are constructed, see Wennberg, "Use of Medicare Claims Data," 2004.

[18] These measures are highly relevant to the current debate concerning the number of physicians that the nation should train. According to the NYU and UCLA benchmarks, there may be a significant deficit; but according the experience of Stanford or UCSF, there may be more than enough. Given the Fisher finding of no marginal benefit with increased care intensity, together with the association between physician supply, utilization, and costs and the lack of consensus among academic medical centers on how to optimally employ the existing workforce, a prudent policy would look for better evidence—based on efficiency and effectiveness—that indicates more are needed.

[19] Variations in volume of care among Blue Cross clients in Michigan have been shown to parallel variations in Medicare. See J. E. Wennberg and D. E. Wennberg, eds., *The Dartmouth Atlas of Health Care in Michigan* (Hanover, N.H.: Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 2000), http://www.bcbsm.org/atlas/.

[20] This recommendation led to Section 646 of the Medicare Modernization Act of 2003, which has yet to be implemented.

APPENDIX. HOW WE MEASURE PERFORMANCE

The essence of practice-variation studies is the comparison of medical care use rates among defined populations. Here is a brief review of how the patient populations have been formulated in the examples used in this report.

- Sometimes the "population at risk" is the whole population living in a region. For example, the incidence of surgery for hip fracture has been measured by counting the number of Medicare residents who had a specific procedure during a given period of time (the numerator of the rate) and dividing by the region's total number of residents who are Medicare enrollees (the denominator). With the exception of lumpectomy, the rates of discretionary surgery discussed in this report are calculated this way, as are a few examples of supply-sensitive care. Typically, such rates are adjusted for differences in age, sex, and race.

- Sometimes the populations selected for comparison are limited to those at the same stage in the course of illness. The denominator for lumpectomy rates is women with early stage breast cancer who had breast cancer surgery. Regional measures of supply-sensitive care at the end of life are based on the medical care received during the patient's final six months. In that case, the denominator is the number of patients who died; and the numerator is the number of pertinent events—for example, days spent in intensive care units—experienced by patients during the last six months of their lives. Because most Medicare enrollees are quite sick during the last six months of life, utilization rates during this period are implicitly adjusted for severity of illness; further adjustments include those for age, sex, race, and, in some examples, possible differences in case mix.

- Sometimes the populations are limited to those with specific illnesses or medical needs. Most measures of the quality of effective care involve such specific populations. For example, in measuring the quality of care for diabetic patients, the numerator is the count of all diabetic patients who received the needed eye examination at least once over a two-year period. The denominator is the count of all diabetic patients living in the region.

- The hospital-specific measures for supply-sensitive care use as the denominator all Medicare enrollees who died from one or more of 12 chronic illnesses.

41

RELATED PUBLICATIONS

Publications listed below can be found on The Commonwealth Fund's Web site at www.cmwf.org.

Medicare Extra: A Comprehensive Benefit Option for Medicare Beneficiaries (October 4, 2005). Karen Davis, Marilyn Moon, Barbara S. Cooper, and Cathy Schoen. *Health Affairs* Web Exclusive (*In the Literature* summary). In this article, a group of health policy analysts propose a new comprehensive benefit option for the traditional fee-for-service Medicare program, featuring less confusion and lower cost than private drug plan and Medigap offerings.

Unmet Long-Term Care Needs of Medicare–Medicaid Dual Eligibles (October 2005). Harriet L. Komisar, Judith Feder, Judith D. Kasper, and Susan Mathieu, Georgetown University Health Policy Institute. This chartpack draws upon information in the *Inquiry* article noted below. In addition, it presents new information of unmet needs for long-term care among dual eligibles.

Unmet Long-Term Care Needs: An Analysis of Medicare–Medicaid Dual Eligibles (Summer 2005). Harriet Komisar, Judith Feder, and Judith Kasper. *Inquiry,* vol. 42, no. 2 (*In the Literature* summary). In a survey of elderly "dual eligibles" living in the community, more than half of those who need assistance with activities of daily living said they do not receive enough help.

Riding the Rollercoaster: The Ups and Downs in Out-of-Pocket Spending Under the Standard Medicare Drug Benefit (July/August 2005). Bruce Stuart, Becky A. Briesacher, Dennis G. Shea et al. *Health Affairs,* vol. 24, no. 4 (*In the Literature* summary). Under Medicare Part D, beneficiaries will incur high average out-of-pocket costs for prescription drugs, and many will face dramatic changes in spending from quarter to quarter, according to this article's authors.

Medicare: Making It a Force for Innovation and Efficiency (July 2005). Jessica Mittler. This issue brief—prepared for the 2005 Commonwealth Fund/John F. Kennedy School of Government Bipartisan Congressional Health Policy Conference—outlines how Medicare could promote innovation and efficiency throughout the health care system.

The Quality of Antipsychotic Drug Prescribing in Nursing Homes (June 13, 2005). Becky Briesacher et al. *Archives of Internal Medicine,* vol. 165, no. 11 (*In the Literature* summary). The authors of this article report that more than half of nursing home residents receiving antipsychotics were given doses that exceeded recommended maximum levels, received duplicative therapy, or had conditions, like memory problems or depression, for which such drugs are considered inappropriate.

Quality of Health Care for Medicare Beneficiaries: A Chartbook (May 2005). Sheila Leatherman and Douglas McCarthy. This chartbook is the first publication of its kind to provide a comprehensive portrait of Medicare's performance on multiple measures of quality, including effectiveness, patient safety, access and timeliness, and capacity to improve.

Prescription Drug Coverage and Seniors: Findings from a 2003 National Survey (April 19, 2005). Dana Gelb Safran, Patricia Neuman, Cathy Schoen et al. *Health Affairs* Web Exclusive (*In the Literature* summary). According to a national survey, four of 10 seniors did not take all the drugs prescribed to them by doctors in the past year, due to cost, side effects, perceived lack of effectiveness, or the belief that they did not need the medication. The survey results showed that prescription drug coverage varies widely, with a large percentage of low-income seniors lacking any kind of coverage.

Impact of the Medicare Prescription Drug Benefit on Home- and Community-Based Services Waiver Programs (April 2005). Charles J. Milligan, Jr., University of Maryland, Baltimore County. With home- and community-based services waiver programs, many low-income, elderly, and disabled adults enrolled in both Medicare and Medicaid can avoid institutionalization and remain in the community. The author of this issue brief says the impending transfer of prescription drug coverage from Medicaid to Medicare may place many "dual eligibles" in jeopardy.

Medicare Advantage: Déjà Vu All Over Again? (December 15, 2004). Brian Biles, Geraldine Dallek, and Lauren Hersch Nicholas. *Health Affairs* Web Exclusive (*In the Literature* summary). Medicare Advantage has the opportunity to learn from the experiences of its predecessor, Medicare+Choice. In this article the researchers consider six challenges facing the program, including simplifying health plan choices for enrollees and addressing plans' efforts to avoid enrolling sicker, higher-cost beneficiaries.

Medicare Disadvantaged and the Search for the Elusive 'Level Playing Field' (December 15, 2004) Robert A. Berenson. *Health Affairs* Web Exclusive (*In the Literature* summary). Medicare Advantage private plans benefit from healthier, less costly beneficiaries and from higher federal payments. But, in the face of growing budget deficits and strong public pressure, will these payments prove sustainable?

Are the 2004 Payment Increases Helping to Stem Medicare Advantage's Benefit Erosion? (December 2004). Lori Achman and Marsha Gold, Mathematica Policy Research, Inc. The MMA provided Medicare Advantage plans with significant increases in monthly payment rates, beginning March 2004. About one-half of the payment increases were used by plans to reduce enrollee premiums and cost-sharing and enhance benefits; providers received most of the rest.

The Cost of Privatization: Extra Payments to Medicare Advantage Plans—2005 Update (December 2004). Brian Biles, Lauren Hersch Nicholas, and Barbara S. Cooper. This issue brief examines the payments that private plans are receiving in 2004 relative to costs in traditional fee-for-service Medicare, using data from the 2004 Medicare Advantage Rate Calculation Data spreadsheet. The authors find that, for 2004, Medicare Advantage payments will average 8.4 percent more than costs in traditional fee-for-service Medicare: $552 for each of the 5 million Medicare enrollees in managed care, for a total of more than $2.75 billion.

Envisioning the Future of Academic Health Centers (February 2003). Offering a blueprint for the future of the nation's teaching hospitals and affiliated medical schools, this final policy report of the Commonwealth Fund Task Force on Academic Health Centers calls for a host of public policy and private management changes intended to strengthen the leadership role of academic health centers and preserve their important research, education, and clinical care missions at a time when they are critical to the nation's well-being.