Variation in Use of Medicare Services Among Regions and Selected Academic Medical Centers:

Is More Better?

Presentation by John E. Wennberg, MD, MPH

Duncan W. Clark Lecture
New York Academy of Medicine

New York, New York
January 24, 2005

On January 24, 2005 I gave the annual Duncan Clark Lecture at the New York Academy of Medicine. I took the opportunity to provide an update on our efforts to better understand practice variations, particularly variation that is unwarranted because it cannot be explained on the basis of illness, patient preferences or the dictates of evidence-based medicine. Don Berwick asked me to prepare this overview of the lecture for the Institute For Healthcare Improvement website <http://www.ihi.org/ihi>. It is also available on the Dartmouth Atlas website <http://www.dartmouthatlas.org/> (Posted 2/24/05.)
INTRODUCTION

By some accounts, health care in the United States has entered a death spiral of ever escalating costs attended by progressive lose of entitlement as more and more employers elect not to provide health insurance and those who do shift the burden of rising health care costs onto their employees. At the same time, Medicare is headed toward a trillion dollar train wreck. Some still hold out the hope that what has become known as pay-for-performance will save the day. Instead of paying all providers at the same rate, reimbursement must be changed to differentially reward providers who show excellent performance in meeting practice guidelines that define “high quality care”. Others believe that the answer rests in making consumers better purchasers of health care by modifying health insurance to include high deductibles and a medical savings account, which, if spent wisely, would lead to a more rational medical market.

Our studies of practice variation hold some good and some bad news for these initiatives. Quality initiatives are now focused on reducing the underuse of effective care such as a beta-blocker drug after a heart attack. While paying providers to do things they should 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 health care dollars, at least for Medicare, go to other categories of care where the quality problem is not underuse.

More than 50% of Medicare spending is used to buy visits to physicians, diagnostic tests and hospitalizations, mostly for patients with chronic illness. Here, (as I will explain) the most important problem is overuse because more is not better, particularly with regard to inpatient care. Unless pay-for-performance focuses on rewarding providers who are efficient in the delivery of these services it will have little impact on overall costs and poor quality associated with too much care.
A significant proportion of Medicare spending is for discretionary surgery. Here, the quality problem is misuse of care, a diagnosis we reach on the basis of evidence that provider opinion rather than informed patient preference is an important driver of utilization rates. As I will elaborate in my lecture, the good news is that the introduction of informed patient choice tends to lead to a decline in demand for higher cost, more invasive treatment options such as discretionary surgery. However, unless pay-for-performance can reward providers who improve the quality of patient decision making for such services, it will have little impact on the cost and quality of preference-sensitive care.

The impact on cost and quality of care of high deductibles and medical savings accounts, we believe, will also depend on the type of care. If targeted to influence patient decisions that involve expensive treatments such as discretionary surgery, making sure that patients have a financial stake may improve the quality of decision making as patients seek information on treatment options and reward providers that offer high decision quality. However, as a tool for addressing the use of care among the chronically ill, the strategy is problematic. The volume and costs of such care becomes progressively higher as illness progresses, reaching a crescendo toward the end of life. Even well endowed savings accounts would be exhausted relatively early in the course of serious chronic illness and thus have little influence on longitudinal costs of managing such patient populations. Moreover, when the goal is to increase appropriate utilization, as it is in the case of immunizations and other examples of effective care, it is hard to see how financial considerations that discourage patient access can help improve quality.

I have three goals for tonight’s lecture. One is to persuade you that our strategy for classifying health care services into effective, preference-sensitive and supply-sensitive care is a useful way of viewing unwarranted variations. This is so because the causes of variation and the normative interpretation of the “right rate” differ according to category, as do the remedies, for example, the design of pay-for-performance or the applicability of medical savings accounts. The second goal is to review recent progress in developing provider-specific performance
measures using Medicare claims data. I illustrate these measures by applying them to an evaluation of how well-known academic medical centers manage chronic illness. Hospital-specific measures are an essential ingredient in pay-for-performance remedies designed to address the overuse of supply-sensitive care. The third goal is to consider briefly the requirements for achieving real and sustainable improvements in quality and efficiency in each of the 3 categories of care.

**EFFECTIVE CARE**

Effective care interventions are viewed as medically necessary care on the basis of clinical outcome evidence, preferably from randomized clinical trials. In this 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 is underused. In a recent publication in the *New England Journal of Medicine*, Beth 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. The researchers found that patients received recommended care only 54.9% 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 cannot be accurately defined in the claims (i.e. the subgroup of heart attack patients needing beta-blockers at discharge) or the item of necessary care is not paid for by Medicare (i.e. drugs such as beta-blockers). However, several can be calculated, and for those we have been able to measure, we have found extensive underuse of effective care. For example, practice guidelines call for an eye examination at least once every two years for people with diabetes (Figure 1). In several hospital referral regions in 2001, fewer than 50% of Medicare enrollees

---

with diabetes had eye examinations; in the “best regions” about 75% of enrollees had them. In regions in and around New York City, rates were above average, but not exemplary. For example, 64.3% of diabetic residents of Manhattan received recommended care, and in the Bronx 63.4% did.

I believe that the underuse of effective care relates in large part to the lack of the infrastructure necessary to support systematic compliance with guidelines. 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 guideline compliance that exceeds that of fee-for-service medicine. It is of interest that 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. Patients with chronic illness who live in high-spending regions also tend to have many more physicians involved in their care, raising questions about who is in charge and responsible for assuring that needed care is provided.²

I have no doubt that finding a remedy for unwarranted variation will depend on finding ways to change the economic incentives that influence all our behaviors, but most specifically the behavior of patients and providers. However, the changes need to be specific to the category of care. For effective care, where physicians are paid to do more, incentives are well aligned and the effects on the health care economy would be minimal as these services themselves cost little. Identifying patients in need will become easier and easier as electronic medical records become more widely used, and the process 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 will of itself lead to a reduction in per capita spending. *To have a significant*

---

² Baicker K, Chandra A. Medicare Spending, The Physician Workforce, And Beneficiaries’ Quality Of Care. Published April 7, 2004 as Web-Exclusive on Health Affairs website: [http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.184v1/DC1](http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.184v1/DC1)
impact on Medicare costs, pay-for-performance strategies must be directed to the other two categories of care.

PREFERENCE-SENSITIVE CARE

Preference-sensitive care comprises treatments that involve significant tradeoffs affecting the patient’s quality and/or length of life. The surgical options for the treatment of early stage breast cancer, for example, usually include mastectomy (complete removal of the breast) or lumpectomy (“breast-sparing surgery,” a local excision of the tumor). The consequences for those women who choose mastectomy include the loss of the breast and, for some, the use of a prosthesis or undergoing reconstructive surgery. For women who choose breast-sparing surgery, the consequences can include radiation and/or chemotherapy and living with the risk of local recurrence, which will require further surgery.

The Dartmouth Atlas has noted striking variations in the proportion of early stage breast cancer patients who undergo lumpectomy. In an early study (1992-93), we found regions in which virtually no Medicare women underwent lumpectomy, but one region in which almost 50% did. Sometimes, adjoining regions had strikingly different rates. For example, in the Elyria, Ohio, hospital referral region, 48% of Medicare women had breast-sparing surgery for early stage breast cancer, while in Cleveland 23% did and in Columbus 12% did.

I have long held to the theory that idiosyncratic practice style is the major source of such variation in rates of discretionary surgery. This theory was first advanced in the 1930s by J. Alison Glover, a British pediatrician, whose studies revealed about a ten-fold variation in tonsillec- tomy rates among school districts. An important aspect of Glover’s findings was that the important decision maker on the need for tonsillectomy was a single physician, the school health officer who routinely examined students for signs of sickness. To the best of his ability given the available data, Glover ruled out the contributions of a number of environmental and illness-related factors to the remarkable variation. His most
convincing evidence, however, was the “natural experiment” that occurred with the advent 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 ten, and remained low for years afterwards. Glover attributed the drop in rates to the change in “medical opinion” embodied in the different practice styles of the two physicians.

Alan Gittlesohn and I, together with two physicians from Morrisville, Vermont, reported a similar ten-fold variation in tonsillectomy rates among Vermont regions in the early 1970s. After the physicians in Morrisville became aware of the high rate in their area, local medical opinion changed radically, and Morrisville rates dropped to near the bottom of the distribution.

A common counter argument to the practice style theory is that patient preferences dominate decision making, and that rates of surgery vary in proportion to variations in preferences. Under this theory, the interpretation would be that 48% of women with early stage cancer in Elyria preferred lumpectomy, while in Columbus only 12% did, and in Rapid City, South Dakota, only 1% did. The question is whether in “usual practice” the physician’s recommended course of treatment corresponds reasonably closely to the patient’s informed preference. Experimental evidence that physicians’ opinions and patients’ preferences for treatment might not be well correlated comes from clinical trials of shared decision making, aided by patient decision aids.

Shared decision making is the process of interacting with the patient to help him or her “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

---

3 O’Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying Unwarranted Variations In Health Care: Shared Decision Making Using Patient Decision Aids. Published October 7, 2004 as Web-Exclusive on Health Affairs website: [http://content.healthaffairs.org/cgi/content/full/hlthaff.var.63/DC2](http://content.healthaffairs.org/cgi/content/full/hlthaff.var.63/DC2)
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 values and preferences; and (3) guidance or coaching in deliberation, designed to improve the patient’s involvement in the decision process.

Clinical trials of patient decision aids have now been completed for a number of conditions that involve 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 due to coronary artery disease; and the choice between surgery and conservative management for patients with back pain due to disc disease. The trials show that, compared to a control group, patients who use decision aids are better informed about the risks, benefits and clinical uncertainties associated with the treatment options available to them. Moreover, the choices patients make in the shared decision making environment (aided by patient decision aids) are “better” decisions: they more closely reflect the patient’s own individual values. Finally, most 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 whose level of illness makes them clinically appropriate candidates for surgical intervention may well exceed the amount of surgery actually being done in a given region. A recent study by Hawker and her colleagues in Canada speaks to this point. The researchers conducted a population-based interview study to screen for patients with pain that might signal arthritis of the knee. The patients so identified were then examined by physicians and given X-ray examinations to define a patient population that would benefit from knee surgery. 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 ten. 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% indicated a preference for surgery; the vast majority wanted conservative treatment.

Let me turn to examining the distribution in rates among hospital referral regions of three orthopedic procedures: knee replacement, hip replacement, and back surgery. All vary remarkably, particularly when compared to hip fracture repair. Knee replacement and hip replacement are four and five times more variable than hip fracture repair, respectively, and back surgery is about seven times more variable (Figure 2). Of interest to a New York audience are the consistently low rates in Manhattan. For example, Manhattan rates of knee replacement ranked second from the bottom among the 306 regions, a full 51% below the national average. Rates of hip replacement and back surgery ranked 273rd and 300th from the top and were 29% and 50% below the national average, respectively.

There are sometimes remarkable differences among neighboring regions. One example is the “surgical signature” of four South Florida communities. Figure 3 compares the rates of surgery in Miami, Fort Lauderdale, Fort Myers and Sarasota to rates in Manhattan. This comparison might be of particular interest because Medicare residents of Manhattan commonly winter in Florida. (New York physicians might want to assess their snowbird patients’ risk of surgery, depending on where they go.) As it turns out, in the years 2000-01, the rate of knee surgery in Fort Myers was three times higher than the rate in 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 in Manhattan, and back surgery rates were more than three times higher in Fort Myers and Sarasota, and two times higher in Fort Lauderdale, than in Manhattan. 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% lower in Miami while the rate of knee surgery was 26% higher and the rate of back surgery was 39% higher.

In theory, the differences among these communities in rates of knee and hip replacement and back surgery could reflect differences in patient preferences about treatment, or the incidence of osteoarthritis and/or herniated discs. In light of the evidence, this seems unlikely. Moreover, there is no epidemiologic evidence that illness or informed patient preferences vary as sharply according to the boundaries of health care markets as does surgery. It seems very unlikely that differences in illness rates and/or patient preferences 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, I believe, rests in the propensity of local surgeons to specialize in a particular subset of the orthopedic surgical workload and in the workforce’s ability to find candidates that meet clinical appropriateness criteria. In the case of knee and hip replacement, the surgical rates in most regions are generated by clinical decisions made by small groups of orthopedic surgeons. Orthopedic surgeons have many options regarding the clinical conditions in which they can subspecialize, including trauma, sports medicine, carpal tunnel syndrome and knee, hip and back conditions. 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 United States average. (A full exploration of these ideas requires data covering the entire resident population, not just Medicare patients.)

An examination of the association between the per capita supply of surgical specialists and the rates of procedures that that 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, knee and back 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 R² values of .01 and .06, respectively.) The relationship between the supply of orthopedic surgeons and rates of back surgery has an R² value of .02.5

The persistence of surgical signatures over long periods of time supports the interpretation I have already suggested that surgical specialists tend to become expert in a subset of the procedures that their specialty performs, and to orient their workload toward patients eligible for the procedure they are comfortable with (Figure 4). The figure shows the surgical signatures of the Fort Myers, Fort Lauderdale and Miami hospital referral regions over a decade, and the rates relative to the rate in Manhattan each year. Note the year-in, year-out consistency in the rates. Over the decade, the differences in rates add up to substantial differences in the number of procedures performed. For example, the surgeons working in Fort Myers performed 7,250 more back operations, 7,000 more knee replacements, and 2,600 more hip replacements than would have been done had the Manhattan rate prevailed in those communities. Compared to Manhattan, over the ten-year period Miami “produced” 870 more back surgeries and 1,400 more knee surgeries — but 56 fewer hip replacements.

5 The absence of a strong association between the per capita supply of orthopedic surgeons and rates of knee and hip replacement and back surgery is also true of other surgical specialists and 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 (check this out), there is little association between the per capita supply and the rates of common procedures performed by those specialists. The R² statistic ranged from .00 for the association between urologists per capita and transurethral prostatectomy to .09 for the association between vascular surgeons and lower extremity bypass grafting.
The stability of the surgical signature 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 1992-1993 and 2000-2001 is 0.75 – that is, 75% of the variation in rates of knee surgery in 2000-01 is “explained” by the rate in 1992-93. Interestingly, in the vast majority of hospital referral regions, rates in 2000-2001 were higher than they were in 1992-1993. The United States average rate increased by 40%. The supply of orthopedic surgeons increased about 9%. Even though both the supply of orthopedic surgeons and the overall rates of surgery increased, 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?

Under the normative assumption that the “right rate” of a given procedure should be based on the choices made by informed patients (free of undue influence from the practice styles of their physicians or other inappropriate influences), the systematic implementation of shared decision making, supported by decision aids, would offer the opportunity to establish valid measures of the actual demand for a given treatment option. Such an opportunity presented itself in the early 1990s. A decision aid designed to help patients decide between watchful waiting and surgery for their enlarged prostates was introduced in the urologic clinics in two pre-paid 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%, 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 that resulted from shared decision making was at the extreme low end of the national distribution, suggesting that the rates of prostate
surgery in most regions of the United States might substantially exceed the amount that informed patients actually want.\textsuperscript{6}

The results of clinical trials of decision aids and observational studies of their impact on population-based rates suggest that the amount of discretionary surgery performed in the United States exceeds the amount that informed patients want. It is not clear, however, what the steady state demand for discretionary surgery would be over time if shared decision making were fully implemented in primary care as well as specialty practice. Many patients who might want surgery may escape referral because of the practice styles of their primary physicians. Moreover, patients’ preferences about discretionary procedures aimed at improving their quality of life, such as knee and hip replacement, might change over time as their conditions progress, becoming more painful or function-limiting. We know relatively little about these possibilities, since shared decision making supported by patient decision aids has not yet been systematically implemented, even in pre-paid group practices such as Kaiser Permanente.

What it is safe to conclude, however, is that current patterns of practice do not reflect demand based on patient preferences, and that geographic variations in rates of surgery that reflect physician practice style will persist until patients are actively involved in the decision process and there are incentives for physicians to adopt shared decision making.

The economic impact on a health care market of the introduction of shared decision making for preference-sensitive care involving discretionary surgery will depend on the level of spending and the magnitude and direction of its effect on surgery rates. For example, over the ten-year period from 1992-2001, we estimate that Medicare spending (in 2001 dollars) for knee and hip replacement and back surgery in Fort Lauderdale and Fort Myers was $137 million and $135 million.

million, respectively, 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 is thus potentially disruptive to both surgeons and hospitals, even though it results in cost savings for the payers and better quality of care for patients. If shared decision making were implemented at the hospital level, pay-for-performance strategies would presumably require substantial sharing of savings as part of the incentive package for the providers, a strategy which, ironically, could end up disproportionately rewarding regions and hospitals with high rates. An alternative might to be to construct the pay-for-performance intervention at the primary care level.

The likelihood of health savings accounts (HSAs) having a significant impact on Medicare program cost also varies, I believe, according to the category of care. For preference-sensitive care, HSAs could conceivably lead to better decisions. The introduction of economic considerations into decisions that involve tradeoffs that depend on patient values would presumably have an additional impact on patient choice, particularly in situations where one treatment is much more costly than another.

SUPPLY-SENSITIVE CARE

The third category of care is supply-sensitive care. Most supply-sensitive care (which appears to account for at least 50% of medical spending) is provided to people with chronic illnesses. The conditions that generate the most spending are 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. 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).

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. For patients at a given stage in the progression of chronic illness, medical textbooks contain no evidence-based clinical guidelines for scheduling patients for return visits, when to hospitalize or admit to intensive care, when to refer to a medical specialist, and, for most conditions, when to order a diagnostic or imaging test. As an example, the pages of 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” -- contain not a single reference as to when to hospitalize or schedule for a revisit patients with cancer, chronic lung disease or heart failure.7

In the absence of evidence and under the generally held assumption that medical resources should be fully utilized in the management of seriously ill patients, it should not be surprising that the supply of resources governs the frequency of their use. Over the years, the Dartmouth Atlas Project has consistently shown a positive association between the supply of staffed hospital beds per 1,000 residents and the hospitalization rate for medical (non-surgical) conditions (Figure 6). More than half of the variation in discharge rates is associated with bed capacity. By contrast, hospitalizations for hip fracture -- one of the few conditions in which variation closely reflects the incidence of illness -- has little correlation with resource supply.

The effect of hospital bed supply on hospital use has long been recognized, gaining in the 1960s a status near that of a law of nature; it has in fact become

---

known as “Roemer’s law.” In looking at the association between hospital beds and utilization, it is important to remember that the effect is primarily on patients with medical (i.e. non-surgical) conditions. Hospitalizations for major surgery — whether in the preference-sensitive or effective care categories — are not correlated with overall beds per capita. In my 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. I gained this impression from interviews with clinicians practicing in Boston and New Haven, who were not aware of the 60% 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.

The use of hospitals for the treatment of people with medical conditions is particularly intense during the last few months of life, and the variation among regions is striking. On average, patients living in the lowest rate regions spent about six days in hospitals, while those in the highest rate region spent twenty days (Figure 7). It is of interest that the hospitalization rate among Medicare residents of Manhattan was the highest in the nation. Since by definition all enrollees were dead by the end of the period, we do not believe that confounding due to unmeasured differences in illness played an important role in the variation.

A similar relationship can be seen between the supply of physicians and visit rates, particularly for those specialties that spend most of their time treating chronic illness. Figure 8 illustrates the relationship between the number of cardiologists per 100,000 residents of the region and the number of visits to cardiologists per Medicare enrollee. About half of the variation in the number of visits to cardiologists per Medicare enrollee is associated with the number of

---

8 In honor of Milton Roemer, see reference from The Dartmouth Atlas of Health Care: In the early 1960s, Milton 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 provides an important illustration of what has become known as “Roemer’s Law.”
cardiologists per 100,000 residents. A similar relationship exists between the supply of internists and numbers of visits to internists.

The behavioral basis for this association seems be that the Medicare population constitutes a large share of the cardiologists’ and internists’ patient loads. Appointments to see physicians characteristically are fully “booked” -- very few hours in the work week go unfilled. Most office visits are for established patients, and the interval between revisits is dictated by the size of the physician’s panel of patients. On average, regions with twice as many cardiologists per 100,000 residents will have twice as many available office visit hours. In the absence of evidence-based guidelines on the appropriate interval between revisits, available capacity governs the frequency of revisits.

Physician visit rates among people who are in their last six months of life vary substantially. In the highest-rate region, these people 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). The East Long Island and Manhattan hospital referral regions are near the top of the national distribution.

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 question has received virtually no attention from academic medicine or federal agencies, such as the National Institutes of Health, which 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 haven’t been done. The issue of the appropriate quantity of supply-sensitive care is only beginning to emerge as a topic for medical discourse at medical rounds and in scientific journals and textbooks.
A recent population-level study by Elliott Fisher and colleagues at Dartmouth⁹ provides a provisional answer about whether regions with greater intensity of clinical practice have better outcomes. The researchers examined the outcomes of three patient cohorts enrolled because they had a hip fracture, a heart attack or a colectomy for colon cancer. The study compared cohorts living in regions with greater care intensity to those in regions with less intensity. The measure of care intensity was price-adjusted end-of-life spending. (As I said, we believe end of life measures provide valid measures of relative intensity of care that is untainted by differences in illness.) The question was whether those living in regions where they are likely to receive more care had better outcomes than those living in regions where they were likely to receive less. The patients were followed for up to five years after their initial event – the hip fracture, surgery for colon cancer, or heart attack. The study showed increased mortality rates in regions with greater care intensity.

Figure 10, adapted from the Fisher study, compares the level of resource inputs and mortality among cohorts living in regions in the highest quintile of Medicare end of life spending with those living in the lowest quintile. Relative to the experience of those in the lowest quintile, Medicare spending was 61% higher (on a price-adjusted basis). High-rate regions had 32% more hospital beds per capita, 31% more physicians, 65% more medical specialists, 75% more general internists and 37% more surgeons. The low-rate regions had 25% more family practice physicians than the high-rate regions.

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% higher for hip fracture patients, 1.2% higher for colon cancer patients, and 5.2% higher for heart attack patients.

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

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 their hip fractures, patients using academic medical centers in high-spending areas had 82% more physician visits, 26% more imaging exams, 90% more diagnostic tests and 46% more minor surgery. Compared to low-intensity regions, patients with hip fractures, colon cancer and heart attacks who were loyal to academic medical centers 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 we 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 their illnesses. The study populations were defined by


Wennberg, JE, Fisher, ES, Stukel, TA, Sharp, SM. Use of Medicare claims data to monitor provider-specific performance among patients with severe chronic illness. Published October 7, 2004 as Web-Exclusive on Health Affairs website: http://content.healthaffairs.org/cgi/content/abstract/hlthaff.var.5v1
assigning enrollees to the hospital they most frequently used in the two years prior to death. We selected for comparison 77 hospitals that appeared on the 2001 list of the “best” hospitals for geriatric care and for heart disease, cancer and pulmonary disease published by *U.S. News and World Report*. Most of these hospitals are well-known academic medical centers.

We profiled the management styles of these institutions using several measures. The efficiency measures include use of hospitals, ICUs, physician visits and Medicare spending. The quality measures include the proportion of deaths occurring in ICUs and the proportion of the population seeing ten or more physicians during the last six months of life. Resource input variables included FTE physician inputs. Although selected for their reputations for high-quality care, the *U.S. News and World Report* hospitals differed remarkably among themselves in the way they managed severely ill Medicare patients. This was true even among hospitals in the same states and cities.

**Average number of days spent in hospitals** during the last six months of life ranged from 9.4 to 27.1 per decedent (Figure 11).\(^{11}\) Three of the hospitals on the *U.S. News and World Report* list were located in Manhattan, and four were in California. Patients assigned to the three academic medical centers in Manhattan had the highest patient day rates among the 77 hospital cohorts. Patients loyal to the New York University Medical Center spent almost a month in the hospital, while those assigned to Mount Sinai and New York-Presbyterian\(^{12}\) hospitals spent 22.8 and 21.6 days, respectively. There were striking differences in patterns of utilization within California. 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

---

\(^{11}\) Rates are based on all hospitalizations during the last six months of life, the vast majority of which occurred in the hospitals to which the patients were assigned. Severe chronic illness was defined as patients with complicated illness in 12 categories proposed by Iezonni and her colleagues. Rates are adjusted for age, sex, race, and type of chronic illness. For full listing of results (including confidence limits) see [http://www.dartmouthatlas.org](http://www.dartmouthatlas.org).

\(^{12}\) The estimate is the weighted average for the two organizations. Because of the way the hospitals were coded, the experiences of New York Hospital and Presbyterian Hospital could not be examined separately.
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 Medical Center in Los Angeles (UCLA) spent 16.1 days in that hospital, 24% fewer than patients loyal to Cedars-Sinai, but 40% more days than among those loyal to its sister organization, the University of California Medical Center in San Francisco (11.5 days).

We found that hospitals that had 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 hospital for patient cohorts with congestive heart failure (CHF) and cancer were highly correlated ($R^2 = .64$) even though, on average, cancer patients tended to be hospitalized less (Figure 12). There were similar correlations between the rates of hospitalization for chronic lung 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 is the hospital to which the patient was assigned, not whether they had cancer or CHF.

We also analyzed 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 patients have or their relative need (as indicated by ethnicity)? The occupancy rates for the 77 hospitals were uniformly high, meaning that available beds were fully utilized. The behavioral basis, I believe, is the influence that capacity has on the decisions physicians make regarding the need for hospitalization. Patients with CHF, COPD and cancer are quite sick,
particularly during the terminal phases of their illness. Physicians find it easier to manage complex patients in the hospital. 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. A similar threshold effect seems to apply to the race-specific risk of hospitalization. While blacks have slightly higher use rates (perhaps reflecting their relative lack of alternatives to hospital care), the effect on hospitalization rates of the particular hospital to which they are loyal is much stronger than the effect of ethnicity.

The average number of days spent in intensive care units during the last six months of life 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 UC hospital, UCSF (2.6 days). Stanford’s use of ICU beds was 1.6 times greater than UCSF’s. There are equally interesting contrasts in New York. New York University Medical Center patients spent an average 6.7 days in intensive care units, 2.4 times more than patients loyal to Mount Sinai (2.8 days), while New York Presbyterian patients (4.5 days) spent 1.6 times more days in ICU than patients loyal to Mount Sinai Hospital.

As was the case with days in hospital, days in ICUs were highly correlated among patients with different chronic illnesses and ethnicities. Although we do not have measures of the number of ICU beds or their occupancy rates, ICU beds also tend to be fully used, implying that the quantity of beds is not closely calibrated to the size of the population loyal to the hospitals in our study. 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, is unclear.

The average number of 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 made an average of 40.3 visits per decedent during the last six months of life. 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 more visits than the average among patients loyal to Stanford. UCLA visits rates (43.9 per decedent) were 61% higher than UCSF rates, and 93% higher than rates among patients loyal to Stanford, but 34% lower than rates among patients loyal to Cedars-Sinai.

Patients who spend more days in hospitals have more physician visits, as shown by the strong ($R^2 = .60$) association between days spent in hospital and physician visits among the 77 hospitals in the cohort (Figure 16). The basis for this association, I believe, is that referrals and revisits are much more easily scheduled when the patient is in the hospital. As well, 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.

The percent of patients who saw ten or more physicians during the last six months of life varied from less than 17% to more than 58% (Figure 17). Mount Sinai and NYU were at or near the top of the distribution: 58.5% and 57.1% of patients assigned to these hospitals saw ten or more physicians. At New York Presbyterian, the rate was 37.7%. 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% and 48.2%, respectively, saw ten 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 ten 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,” where lots of physicians get involved in care but no one is responsible for the coordination of care. The inverse association between percent of physicians involved in caring for chronically ill patients and
scores on quality measures (i.e., percent in need who get effective care) is consistent with this interpretation.\textsuperscript{13}

The percent of patients who die in intensive care units. 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. We see striking differences among academic medical centers in the chance of dying in an intensive care unit.\textsuperscript{14} 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\% 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. \textit{Greater intensity of terminal care, with its negative impact on the quality of dying, is thus not a price the dying must pay to assure overall greater survival among those with chronic illnesses.} 

\textbf{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 three-fold among hospital referral regions and academic medical centers. Regions and academic medical centers with greater overall spending rates do not have higher quality of care, as measured by effective care measures; nor, perhaps surprisingly, do they have higher rates of discretionary surgery. Greater spending is the result in large measure of the providers in these regions having higher utilization rates for supply-sensitive care: more physician visits, hospitalizations, stays in ICUs, and

\textsuperscript{13} Baicker K, Chandra A. Medicare Spending, The Physician Workforce, And Beneficiaries’ Quality Of Care. Published April 7, 2004 as Web- Exclusive on Health Affairs website: \url{http://content.healthaffairs.org/cgi/content/full/hlthaff.w4.184v1/DC1}

\textsuperscript{14} The measure is an approximation, because exact date of discharge from ICU and date of death were not matched.
diagnostic testing and imaging. In view of the Fisher findings, the problem is not underuse in low-rate regions and hospitals, but 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 the last six months of life and spending for the same patient cohort during previous periods prior to death. Even though spending is much lower in previous periods, the rates are highly correlated. 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 and 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 – 24th months prior to death. However, Medicare program spending varied almost three-fold among the 77 hospitals cohorts, from $11,500 to $37,200 per decedent during the last six months and from $2,200 to $8,100 during the 18th – 24th months prior to death. The spending patterns were very highly correlated ($R^2 = .79$) (Figure 19).

We believe that 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 (S-FTEs) invested in the care of the 67 hospital cohorts. 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 ranged from a low of 1.8 to 15.5 S-FTEs per

---

15 These measures were available for only 67 of the 77 hospitals. For information on how these measures are constructed, see: Wennberg, JE, Fisher, ES, Stukel, TA, Sharp, SM. Use of Medicare claims data to monitor provider-specific performance among patients with severe chronic illness. Published October 7, 2004 as Web-Exclusive on Health Affairs website: [http://content.healthaffairs.org/cgi/content/abstract/hlthaff.var.5v1](http://content.healthaffairs.org/cgi/content/abstract/hlthaff.var.5v1)
1,000 decedents, while inputs of primary care range from 2.4 to 10.4 S-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 S-FTEs for Stanford to 24.6 S-FTEs for NYU, a three fold range in variation. The combined primary care-medical specialist inputs to Cedars-Sinai and Mount Sinai cohorts were 20.7 and 16.4 S-FTEs per 1,000 decedents, respectively. UCLA used 59% more physician labor than UCSF. Note from Figure 20 the wide range in variation in reliance on primary care physicians versus medical specialists. For example, the ratio of medical specialists 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 group and staff model HMOs such as Kaiser Permanente. By making them available for fee-for-service organizations, we hope to stimulate accountability for capacity, which we believe, is an essential component of any strategy to reduce the overuse of supply-sensitive care.16

ACHIEVING REAL AND SUSTAINABLE IMPROVEMENT IN QUALITY AND EFFICIENCY

Over the next few months, the Dartmouth Atlas project plans to make hospital-specific information available for California, with subsequent releases in 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 has called longitudinal efficiency) may stimulate payers to seek to reexamine their provider networks (which traditionally have been based on

---

16 We also believe they are highly relevant to the current debate concerning the number of physicians we should train. According to the NYU and UCLA benchmark we may have significant deficit; but according the experience of Stanford or UCSF, we have more than enough. Given the Fisher finding of no marginal benefit with increased care intensity, the association between physician supply and 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 that more are needed.
unit price, not volume x price) and motivate employers to seek to move 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. It should also assure the profitability of those health plans participating in Medicare Advantage that can make deals to send their patients to physician groups using hospitals with spending levels below the regional average. Ironically, traditional Medicare, unless it too can join in directing patients to efficient providers stands to lose: if commercial payers steer patients away from the high cost providers, the population loyal to such providers will shrink, but available resources will not, resulting in yet higher utilization rates and higher costs for supply-sensitive care and possible worse outcomes among the chronically ill Medicare patients who remain loyal to such providers.

In a previous article in Health Affairs, my colleagues and I argued that simply shifting patients to low cost providers would result in little improvement in quality beyond that achieved by reduction in overuse of supply-sensitive care, primarily in the acute care sector. What was needed was reform in financing that would facilitate investment in under funded aspects of care in the non-acute care sectors and in the resources needed for active chronic disease management. We think the availability of provider-specific estimates for the actuarial costs of care discussed above may provide opportunity for new thinking in the design of “budget neutral” reimbursement strategies such as a “partial capitation” that would provide preferred providers with budgets that help compensate for loss in revenue associated with the reduction in inpatient care. We recommended a pilot demonstration project between CMS and progressive health care organizations.

---


18 Wennberg, JE, Fisher, ES, Skinner JS. Geography and the debate over Medicare reform. Published as Web-Exclusive on Health Affairs website: [http://content.healthaffairs.org/cgi/content/full/hlthaff.w2.96v1/DC1](http://content.healthaffairs.org/cgi/content/full/hlthaff.w2.96v1/DC1)
that share the goal of reducing unwarranted variation in all three categories of care.\textsuperscript{19} As experience is gained and the quality of care improves, additional incentives might be put in place to promote improved population-based management such as rewarding managers who use benchmarks from efficient providers in the recruiting of medical personnel and the construction of facilities. We believe the measures of workforce labor input reviewed above could be used for this purpose.

In the long run the most perplexing 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 high quality, low cost care, strategies to assure that all Medicare patients are served by such hospitals remain elusive. If Medicare was willing and able 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, one which would presumably result in accelerated change throughout U.S. health care markets.

\textsuperscript{19} This recommendation led to Section 646 of the Medicare Modernization Act of 2003, which has yet to be implemented.
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. In this lecture I have reviewed examples of effective care, preference-sensitive care and supply-sensitive care, and discussed the causes of unwarranted variation. We find:

• Underuse of most kinds of effective care, such as the use of beta-blockers for people who have had heart attacks and screening of diabetics for early signs of retinal disease. The causes of underuse include discontinuity of care (worse when more physicians are involved in the care) and lack of infrastructure to assure outreach and the timely use of these services. Pay-for-performance should lead to reduction in underuse.

• Misuse of preference-sensitive care, where treatment options involve significant tradeoffs that should be based on the patient’s own values. The causes 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 that of others. Adjustment of economic incentives to reward adopters of shared decision making could lead to a reduction in unwarranted variation. Medical savings accounts may make patients more involved in active participation in decision making.

• Overuse of supply-sensitive care, 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 profiling performance in managing chronic illness can help identify efficient providers. Pay-for-performance and related strategies to reward efficient providers and pay for infrastructure for managing chronic illness could promote reform.
Appendix A: How we measure performance
The essence of practice variation studies is the comparison of rates of use of medical care among defined populations. Let me provide a brief review of how our patient populations have been formulated in the examples I use in this lecture.

• Sometimes the “population at risk” is the resident population living in a region. For example, the incidence of surgery for hip fracture has been measured by counting the number of residents who had a specific procedure during a given period of time (the numerator of the rate) and dividing by the number of Medicare enrollees living in the same region (the denominator). With the exception of lumpectomy, the rates of discretionary surgery discussed in this lecture were 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. Regional measures of supply-sensitive care at the end of life are based on the medical care received during the last six months of life. The denominator is the number of patients who died; the numerator is the number of events experienced by patients during the last six months of their lives — for example, days spent in intensive care units. 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, for 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 are for 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.
Figure 1
Percent of diabetic Medicare enrollees receiving eye exam among 306 hospital referral regions (2001)

Each dot represents the score of a region on a quality measure for diabetic care. The numerator is the number of diabetics who received the medically necessary care, an annual eye examination. The denominator is the number of diabetics living in the region. The figure highlights the location of the hospital referral regions within New York.

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 of four orthopedic procedures: hip fracture repair; knee and hip replacement; and back surgery. Each dot represents one of the 306 regions. The rates are expressed as the ratio to the US average (plotted on a log scale). The numbers in parentheses are the systematic components of variation, measures which allow comparisons of variation among procedures with different mean rates. Knee replacement is about four times more variable than hip fracture repair; back surgery is almost seven times more variable. The numerator is the number of patients with the indicated procedure; the denominator is number of enrollees in traditional Medicare living in the regions.
This figure profiles the rates of knee replacement, hip replacement and back surgery among four South Florida medical communities. The rates are expressed as a ratio to the rate of Manhattan. For example, compared to Manhattan, the rate of knee replacement in Fort Myers is 3.04 times greater. The rates are age, sex and race adjusted.

This table profiles the rates in two year periods over the decade from 1992 to 2001 for knee replacement, hip replacement and back surgery among three South Florida medical communities. Rates are expressed as a ratio to the rate of Manhattan during the corresponding period. Ratios are quite consistent from year to year. Accumulating over the decade, the number of cases in excess of the amount predicted by the Manhattan rates for all three procedures reaches 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 for selected supply-sensitive services. Each dot represents one of the 306 regions. The rates are expressed as the average to the US average (plotted on a log scale). Rates are age, sex and race adjusted. The number in parenthesis is the coefficient of variation. Primary care visits vary about three-fold and demonstrate the least variation. Medical specialists and hospitalizations for COPD vary more than five-fold and CHF discharges about four-fold.

Figure 6
Association between hospital beds per 1,000 (1996) and discharges per 1,000 (1995-96) among Medicare enrollees in 306 HRRs

This figure shows the association between supply of hospital beds per 1,000 residents and the hospitalization rate for medical (non-surgical) conditions. More than half of the variation in discharge rates are 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.
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 hospital by resident Medicare enrollees during the last six months of their life. The numerator is the number of days within six months of death; the denominator is all within the region who died.
Figure 8
Association between cardiologists and visits per person to cardiologists among Medicare enrollees (1996): 306 regions

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)

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. The numerator is the number of visits to physicians within six months of death; the denominator is those within the region who died.
This slide, adapted from the Fisher study, compares on the left the level of resource inputs and on the right the health care outcomes. It compares the regions in highest quintile of Medicare spending with those in the lowest quintile relative to the experience of the population living in the lowest quintile. See text for explanation.


Each dot represents one of the 77 hospital cohorts, presenting the average number of days spent in hospital per person during the last six months of life for each cohort. The figure highlights the patient day rates in the last six months of life for patients assigned to those academic medical centers located in Manhattan and in California.
Figure 12
Association between hospital days for cancer and for CHF patients during last six months of life among 77 “best” U.S. hospitals

The figure examines the relationship between average number of days in hospital for patients cohorts with congestive heart failure (CHF) and solid tissue cancers. Each dot represents the rate for patients assigned to a given hospital.

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

This figure examines utilization rates among black (vertical axis) and white (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 last six months of life among
patients receiving most of their care in one of 77 “best” US hospitals

Each dot represents one of the 77 hospital cohorts, presenting the average number of days spent
in intensive care per person during the last six months of life for each cohort.

Figure 15
Average number of physician visits during last 6 months of life
among patients receiving most care in one of 77 “best” US hospitals

Each dot represents one of the 77 hospital cohorts, presenting the average number of physician
visits per person during the last six months of life for each cohort.
This figure gives the association during the last six months of life between days spent in hospital per person and physician visit per person among the 77 “best” hospitals’ cohorts.

Each dot represents one of the 77 hospital cohorts, presenting the percent who, during the last six months of life, saw ten or more physicians. For example, 59% of the patients who were assigned to Mount Sinai hospital saw ten or more physician visits while only 23% assigned to Stanford did.
Each dot represents one of the 77 hospital cohorts, presenting the percent of deaths associated with hospitalization in an intensive care unit. The numerator is all patients who were admitted to an ICU during a hospitalization that ended in death. The denominator is all members in the cohort and thus includes those who died out of hospital.

This figure correlates Medicare spending (Part A inpatient and Part B) per decedent during the last six months of life and during the period 18-24 months prior to death among the 77 hospital cohorts.
The figure provides estimates of standardized full-time equivalent (S-FTE) labor inputs for primary care physicians (20a) and for medical specialists (20b) among 67 “best” hospital cohorts. See text for explanation.