

Geographical Variations in Medicare Spending

EDITOR'S NOTE

In this issue, we have taken the unusual step of asking three experts to comment on the two lead articles. We did so because we thought that these articles contribute new and important information to our national dialogue about one of the great dilemmas of American life: the high cost of medical care. The articles provide insight into the reasons for variations in costs of care in different parts of the United States and help us to understand what we gain from greater expenditures. These articles are pertinent for us as internists because we decide whether to order many of the services whose frequent use distinguishes higher-spending regions from lower-spending regions.

We asked Kenneth Shine, MD, to discuss the implications of the study findings for physicians. Dr. Shine is a cardiologist. He recently completed 10 years as the president of the Institute of Medicine of the National Academies. He now works on health security issues.

Charles Phelps, PhD, discusses the issues from the perspective of a health economist. He helps us to understand the strengths and the limitations of the methods used in the two articles. Dr. Phelps is the provost of the University of Rochester in Rochester, New York.

We asked Gail Wilensky, PhD, to discuss the articles from the perspective of a Medicare administrator and health policymaker. She was the administrator of the Health Care Financing Administration (now the Center for Medicare & Medicaid Services) from 1990 to 1992. More recently, she chaired two expert panels that advise the U.S. Congress, the Medicare Payment Advisory Commission, and the Physician Payment Review Commission.

We expect these articles to generate a vigorous discussion, and we're grateful to our three editorialists for setting the stage.

Harold C. Sox, MD
Editor

Location, location, location. As in real estate, what you pay for health care varies significantly by where you make the purchase. But there is no guarantee that the quality of the product will vary with price.

In this issue, Fisher and colleagues (1, 2) have provided compelling evidence that 5-year mortality rates, functional status, and quality of care for three conditions (acute myocardial infarction, hip fracture, and colorectal cancer) do not vary significantly from high-cost to low-cost hospital referral regions (HRRs). Costs varied primarily by the number of consultations, tests, and hospitalization days rather than by the evidence-based services required. If anything, mortality rates were somewhat greater in the highest-cost areas. There was wide variation in use of intensive

care unit beds, emergency intubations, and feeding tubes during the last 3 years of life. Influenza and pneumococcal immunizations and Papanicolaou smears were performed less frequently in regions with higher expenditure indexes. Patients in areas with higher expenditure indexes were more likely to see medical subspecialists, and those in HRRs with lower expenditure indexes were more likely to see family practitioners. Although the differences are small, the authors point out some evidence that access to care was poorer in higher-expenditure areas.

This study is consistent with evidence that the more hospital beds, physicians, laboratories, and subspecialists are available in a region, the more they will be used (3). Efforts in the United States to distribute physicians and beds more appropriately for the country's needs have met with limited success. However, our professional organizations must initiate efforts to help physicians and patients understand that more is not necessarily better. Where evidence is available for the proper approach to diagnosis and treatment, it should be applied appropriately. For conditions or circumstances in which evidence is not available, it must be collected. Evidence should drive diagnostic testing and treatment behaviors, and physicians must carefully examine the added value for the patient of each costly decision.

We know all too little about how clinical reasoning differs in regions with different health care expenditure patterns. Development of algorithms for treatment of a specific condition by physicians in high- and low-cost HRRs might be very revealing. Explicitly describing the decision tree (or clinical guidelines) would provide a mechanism for comparing how decisions vary from high- to low-expenditure areas. For example, do the indications for a procedure used in the care of a patient with myocardial infarction differ in different settings? The opportunity to observe the decision-making process for a number of standardized patients in various parts of the United States might characterize the process by which these remarkable variations come about.

The Institute of Medicine report *Crossing the Quality Chasm* (4) emphasized the need for "continuous healing relationships." A physician providing such a relationship has a responsibility to see that the patient receives appropriate evidence-based care but does not receive care that adds no value. The physician can advise on the number and nature of the subspecialists consulted and the frequency of examinations, tests, and procedures. The electronic medical record can have an important role in reducing redundant and unnecessary tests. The patient's continuing-care physician often is aware of which subspecialists practice cost-effectively. This information should influence referral recommendations. General internists and

family physicians can integrate the care process and provide preventive services, which, as Fisher and colleagues show, are not as well provided in areas with large numbers of subspecialists. These physicians can play a crucial role in the appropriate use of intensive care unit beds and procedures at the end of life. These responsibilities emphasize further the importance of the involvement of a generalist physician in the care of every patient.

Clues in the work of Fisher and colleagues suggest that poorer access to care may increase costs. Patients in the most expensive HRRs had the lowest “global” satisfaction and the highest interpersonal satisfaction with care. This suggests that a patient who went to the doctor in the most expensive areas had a more meaningful interpersonal experience but that there was overall discomfort with a system that has large numbers of subspecialists, teaching hospitals, and so on. Reliance on emergency departments rather than on a continuing-care physician may delay the initiation of care. If, in fact, access is decreased in more expensive areas, it may lead to higher costs for care of patients coming later to care—and perhaps a slightly higher mortality rate. May delayed access result in increased services? This hypothesis deserves further exploration.

As Fisher and colleagues point out, if all 306 HRRs had the same costs as the lowest-cost HRRs, as much as 30% of health care costs might be eliminated without adversely affecting health care outcomes. We should test this hypothesis rather than assume that it is true. In the absence of professional leadership, and in the face of rising health care costs, insurers and policymakers could respond with increased co-payments for visits and nonpayment for tests and procedures that are not part of the evidence-based requirements. It would be far better for physicians to reduce the rate at which health care costs are increasing through more efficient care than to invite the blunt bludgeons of purchasers and insurers. Medical educators can

What's Enough, What's Too Much?

Every clinician-scientist admires the randomized, controlled trial (RCT) as the best-known method for studying the effects of a treatment on specified outcomes. A good RCT eliminates (by random assignment) any statistical linkages between the characteristics of the participants (for example, age, unmeasured comorbid conditions) and outcomes. Alas, we have no RCTs for many medical interventions. However, several valuable alternatives have emerged, including case-control analyses and analysis of regional variations in treatment patterns and medical outcomes (“small-area variations”). In this issue, Fisher and colleagues (1, 2) use small-area variation analysis to illuminate the effects of medical resource use on patients presenting with hip fracture, colorectal cancer, and acute myocardial infarction.

The authors characterized a region's propensity to use

work closely with students and house officers to define and redefine the need for visits, consultations, tests, and procedures. The challenge comes when these trainees see their own role models, who are largely subspecialists, maximizing the flow of patients and income. Do patients in such situations have a “continuous healing relationship”?

Fisher and colleagues have convincingly demonstrated that excellent outcomes for patients can be achieved in regions that do less, but do it right. The challenge is to convince the public that this is not about rationing but about better care. The last time the profession ignored rising health care costs, we got “managed care.” Fisher and colleagues offer some clues that if carefully explored and elaborated might allow physicians to take a leadership role in controlling costs rather than wait for the next “market solution.”

Kenneth I. Shine, MD

RAND

Arlington, VA 22202-5050

Requests for Single Reprints: Kenneth I. Shine, MD, RAND Corp., 1200 South Hayes Street, Arlington, VA 22202-5050.

Ann Intern Med. 2003;138:347-348.

References

1. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder ÉL. The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Ann Intern Med.* 2003;138:273-87.
2. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder ÉL. The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. *Ann Intern Med.* 2003;138:288-98.
3. Wennberg JE, Fisher ES, Skinner J. Geography and the debate over Medicare reform. *Health Aff (Millwood)*. 13 February 2002. Available at www.healthaffairs.org/WebExclusives/Wennberg_Web_Excl_021302.htm.
4. Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: Institute of Medicine-National Academy Pr; 2001.

medical resources based on the overall spending patterns in Medicare patients' last half-year of life. These expenditure patterns depended mostly on inpatient treatment choices and specialist and subspecialist use, usually involving discretionary care, in contrast with relatively uniform patterns for “evidence-based medicine” interventions. This supports the premise that regional variations arise from disagreements about proper uses of medical interventions (3, 4).

Most important, the second of the two articles by Fisher and colleagues (2) demonstrates that patients in high-spending regions had no better survival (if anything, slightly worse) than those in lower-spending regions. Separate analyses of a sample of all Medicare beneficiaries further showed no gains in overall patient mortality, functional status, or patient satisfaction.

Small-area variation works best when geographic loca-

tion acts as a proxy for an RCT-like randomization, eliminating links between patients' conditions and the therapy they receive, an approach often called *natural randomization* (5). But various problems can confound natural randomization, including migration (patients moving to where the medical treatment matches their medical conditions or preferences) and medical providers "following" unusually large concentrations of people with particular conditions (for example, cardiologists following elderly people to Florida).

Fisher and colleagues took several steps to ensure that they had achieved natural randomization. For example, they separately analyzed subsets of their sample who had moved from low-intensity regions, those who had moved from high-intensity regions, and those who had not recently moved. They found the same outcomes in each subgroup. They checked for urban–rural differences, for differences in penetration of health maintenance organizations in the regions, for teaching status of hospitals in a region (often a proxy for more complicated patients), and for hospital volume (also related to patient severity and complexity). In no case did their results differ across subgroups, increasing our confidence that hidden or unmeasured patient preferences or medical complexity did not drive their results. They also used information about each study patient's clinical conditions to assess whether underlying complications or complexities affected their results (they did not).

It is important to understand that Fisher and colleagues have not studied the effects of particular medical interventions on outcomes of care for patients with acute myocardial infarction, colorectal cancer, and hip fracture. Rather, they studied the outcomes of care for patients who lived in areas characterized by different medical styles and overall treatment intensities. Thus, their results do not show whether any specific treatment protocols are more or less effective. Rather, their study illuminates the question of whether reducing intensity of medical treatment overall would likely have important effects on patient outcomes. If their natural randomization is effective, then one can infer that patients in the areas with greater intensity gained little or no additional health (or, by Fisher and colleagues' estimates, had actually slightly worse health outcomes), at least for the index diseases studied. Of course, if the small-area variation approach did not achieve natural randomization, then drawing that inference would risk a possible ecologic fallacy.

This question sits at the center of health policy debates: Can we afford to reduce the overall intensity of medical treatment without paying an overly large price in terms of reduced health? Fisher and colleagues show a world in which the productivity curve in medical care has become flat; no more health emerges when treatment occurs at a higher-than-average intensity. This is not to say that improved health care has not provided important gains for the American people, only that increasing intensity at the

upper end of the intensity spectrum is unlikely to improve health outcomes, and indeed (again, subject to the proper caveats), resource use in those areas could be reduced with no decrement in health outcomes.

Obvious questions remain, even if the results of this study are accepted. Do important benefits remain unmeasured? Some studies suggest that using mortality alone may mask some valuable benefits. For example, regional comparisons of patients with heart disease showed no greater survival in Ontario, Canada, where coronary artery bypass grafting and angioplasty are performed infrequently, than in upstate New York, where these procedures are very common; however, the New York patients gained important improvements in other health-related areas, such as chest pain and shortness of breath (6). Fisher and colleagues also measured patient satisfaction (finding no differences across regions), some components of which should correlate with unmeasured quality-of-life variables.

This study's national scope, its careful control of possibly confounding variables, and its underlying design make it possibly the most compelling yet showing that increased treatment intensity does not bring with it commensurate gains in health. Our health care market has many features guaranteed to increase treatment intensity. This study adds greatly to the value of asking, "What's enough, what's too much?"

Charles E. Phelps, PhD
University of Rochester
Rochester, NY 14627

Requests for Single Reprints: Charles E. Phelps, PhD, 200 Wallis Hall, University of Rochester, Rochester, NY 14627; e-mail, charles.phelps@rochester.edu.

Ann Intern Med. 2003;138:348-349.

References

1. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder ÉL. The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Ann Intern Med.* 2003;138:273-87.
2. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder ÉL. The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. *Ann Intern Med.* 2003;138:288-98.
3. Phelps CE. Diffusion of information in medical care. *J Econ Perspect.* 1992;6:23-42. [PMID: 10128077]
4. Phelps CE. Information diffusion and best-practice adoption. In: Newhouse JP, Culyer AJ, eds. *Handbook of Health Economics.* Amsterdam: Elsevier Science BV; 2000.
5. McClellan M, McNeil BJ, Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA.* 1994;272:859-66. [PMID: 8078163]
6. Mark DB, Naylor CD, Hlatky MA, Califf RM, Topol EJ, Granger CB, et al. Use of medical resources and quality of life after acute myocardial infarction in Canada and the United States. *N Engl J Med.* 1994;331:1130-5. [PMID: 7935638]

© 2003 American College of Physicians–American Society of Internal Medicine

The Implications of Regional Variations in Medicare—What Does It Mean for Medicare?

The recognition that Medicare spends substantially different amounts of money on seniors in different regions of the United States is neither a new issue nor an unknown issue. The ability to devise appropriate policy responses to these so-called geographic variations in spending has been another matter.

WHAT'S THE PROBLEM?

Geographic variations in spending are associated with two different problems for Medicare. First, these variations cause payments for managed care plans to differ substantially across the United States. The resulting perceived inequities resulted in the Balanced Budget Act of 1997, legislation that has had its own undesirable consequences. Second, the pressure to restrain Medicare spending, which has been a part of the policy process for the past decade, can be expected to increase dramatically over the coming two decades because of the impending retirement of 77 million baby boomers. To the extent that expenditures in high-spending areas are in part unnecessary or inappropriate, reducing spending in these areas could represent a source of new savings for Medicare.

Before the Balanced Budget Act of 1997, unadjusted Medicare spending at the county level was observed to differ by a factor of 3.5. Actually, this observation showed that premium payments made to health maintenance organizations on behalf of seniors who chose to join one differed substantially depending on where they lived. This meant that seniors in different geographic areas would have access to substantially different benefits paid for by Medicare. Both Congress and seniors who lived in low-payment areas viewed this as inequitable. The 1997 legislation introduced a floor on payments to Medicare managed care plans.

However, Congress ignored the cause of the variation in premium payments: the variation in spending in the underlying fee-for-service program. By reducing the variation in payments only to managed care plans and ignoring the underlying variation in spending on traditional Medicare, Congress reduced geographic inequities for managed care participants but increased market distortions in the choice between traditional Medicare and managed care plans. This happened because health care is delivered in local markets and spending by Medicare now differs substantially in some local markets, depending on whether the senior participates in traditional Medicare or in a managed care plan.

WHY HAS IT BEEN SO HARD TO FIX?

There is widespread agreement that some of the variation in spending is appropriate because it is associated

with differences in the underlying health status of individuals and with differences in input prices. The more difficult issues are associated with residual variations and whether higher-spending areas have “better” outcomes, however defined. The development of effective strategies that might drive high-spending areas to lower spending levels requires a better understanding of why these variations in spending occur and what changing the level of spending is likely to mean for care provided to seniors.

CONTRIBUTION OF THE ARTICLES BY FISHER AND COLLEAGUES

The study reported in this issue by Fisher and colleagues (1, 2) is a careful attempt to assess the types of expenditures associated with regional variations after adjusting for illness levels and prices of medical services. Although I will leave it to Dr. Phelps (3) to assess the adequacy of their methods, by limiting the analysis to patients with three conditions and classifying regions on the basis of spending in the last 6 months of life, Fisher and colleagues were better able to control for many of the confounding factors that have plagued other studies. This makes their findings more credible, although not necessarily relevant to all variations in Medicare spending. The most important finding is that greater expenditures appear to be associated with discretionary services but not with improvements in health outcomes. By including measures that encompass quality of life (such as patient satisfaction and functional status) as well as quantity of life (that is, mortality rates), Fisher and colleagues begin to address some of the most important qualifications that had been associated with earlier studies.

WHAT THE FINDINGS COULD MEAN FOR MEDICARE AND THE CENTERS FOR MEDICARE & MEDICAID SERVICES

Fisher and colleagues' findings, if upheld in analyses that include other medical conditions, would provide the best rationale to date for Medicare and the Centers for Medicare & Medicaid Services, the agency that administers Medicare, to drive down spending in high-expenditure areas of the United States.

The problem is that many of the obvious policies are exceedingly blunt as policy tools. The easiest way to drive down spending in high-spending areas is to use local service use or payment targets. This would link payment rates to overall use or spending in the county, much as now happens with physician payments at the national level and changes in the gross national product. The problem is that this strategy would penalize the most conservatively prac-

ticing physicians because of the practices of the most aggressive physicians in the area—hardly a desirable set of incentives for reducing expenditures. A second strategy would be to enforce the use of national benchmarks for desirable practice patterns, or at least to require justification for any deviation from such benchmarks. Aside from the problem of not having many of the benchmarks that would be needed, this strategy raises questions about the proper role of Medicare in setting standards for the practice of medicine, particularly since the federal government has historically deferred to local decision making in this area (4).

CONCLUSION

An increasing amount of attention has been given to the importance of improving quality and patient safety in Medicare. Medicare now pays for the correction of medical errors but does not pay for the costs of putting systems in place that might prevent these errors. We need to find ways to encourage better practices, not discourage them by creating disincentives. We need more thought about how to reward physicians who practice high-quality conservative medicine. Today is none too soon to begin.

Gail R. Wilensky, PhD
Project HOPE
Bethesda, MD 20814

Requests for Single Reprints: Gail R. Wilensky, PhD, Project HOPE, Suite 600, 7500 Old Georgetown Road, Bethesda, MD 20814-6133.

Ann Intern Med. 2003;138:350-351.

References

1. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder ÉL. The implications of regional variations in Medicare spending. Part 1: The content, quality, and accessibility of care. *Ann Intern Med.* 2003;138:273-87.
2. Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder ÉL. The implications of regional variations in Medicare spending. Part 2: Health outcomes and satisfaction with care. *Ann Intern Med.* 2003;138:288-98.
3. Phelps CE. What's enough, what's too much? [Editorial] *Ann Intern Med.* 2003;138:348-9.
4. Medicare Payment Advisory Commission Report to Congress: Medicare Payment Policy. Washington, DC: Medicare Payment Advisory Commission; March 2001:115.

© 2003 American College of Physicians—American Society of Internal Medicine