



CURRENT APPROACHES TO IMPROVING THE VALUE OF CARE: A PHYSICIAN'S PERSPECTIVE

Howard B. Beckman, Thomas Mahoney, and Robert A. Greene
Rochester Individual Practice Association

November 2007

ABSTRACT: Payers are seeking new ways to bring rapidly increasing health care costs under control. In the past five years, health plans have implemented a variety of programs, the most visible of which have been pay-for-performance, public reporting (PR), consumer-directed health plans (CDHP), and tiering. This paper reviews the emerging concerns about measurement programs generally and specific assessments of the above four types in particular. The authors express considerable concern about the utility of programs focused on judgment (PR, CDHP, and tiering). As an alternative, they encourage the use of programs focused on quality improvement. Judgment-based programs undermine collaboration among stakeholders, making it more difficult to implement the system-wide changes that are needed to significantly improve the value of care. Quality-improvement approaches, tied to incentives and accountability, offer a more constructive model for an effective and efficient health care system.

Support for this research was provided by The Commonwealth Fund. The views presented here are those of the authors and not necessarily those of The Commonwealth Fund or its directors, officers, or staff. This report and other Fund publications are available online at www.commonwealthfund.org. To learn more about new publications when they become available, visit the Fund's Web site and [register to receive e-mail alerts](#). Commonwealth Fund pub. no. 1077.

CONTENTS

| | |
|---|----|
| About the Authors | iv |
| Executive Summary | v |
| Introduction | 1 |
| Evaluating Current Methods to Improve the Value of Care | 3 |
| General Measurement Issues..... | 3 |
| Providing Incentives for Physicians to Change | 9 |
| Pay-for-Performance..... | 9 |
| Public Reporting | 13 |
| Consumer-Directed Health Plans..... | 16 |
| Tiered and Limited Networks..... | 20 |
| Conclusions | 24 |
| Recommendations..... | 28 |
| Notes..... | 31 |

LIST OF FIGURES

| | |
|---|----|
| Figure 1 Higher Stakes Require Higher Rigor | 4 |
| Figure 2 The Win-Lose Cycle | 26 |

ABOUT THE AUTHORS

Howard B. Beckman, M.D., is medical director of the Rochester Individual Practice Association (RIPA)—a 3,200-practitioner community-wide physician organization—which has contracted both for capitated care and consulting services since 1998. Dr. Beckman is also a clinical professor of medicine and family medicine at the University of Rochester School of Medicine and Dentistry, where he conducts health services research and maintains a small internal-medicine and geriatrics practice. Dr. Beckman received his undergraduate degree from Brandeis University and his M.D. from the Wayne State University School of Medicine. His primary administrative and research interests are creating and evaluating effective partnerships. Dr. Beckman served as a co-investigator for the Excellus/RIPA Robert Wood Johnson Rewarding Results grant. This work led to publications under his co-authorship that address the effects of pay-for-performance on reducing overuse, underuse, and misuse of services. Recently published work in the *Journal of Healthcare Management*, for example, demonstrated a positive return on investment (ROI) for the pay-for-performance initiative and improved quality measures in treating diabetes and coronary disease.

Thomas Mahoney, M.D., has been chief executive officer and executive director of RIPA since 2003, following 10 years of service as president of the organization. Dr. Mahoney is a clinical assistant professor of medicine at the University of Rochester School of Medicine and Dentistry and maintains a private practice in general internal medicine. He received his M.D. from the State University of New York at Buffalo School of Medicine. Dr. Mahoney currently serves on the boards of Lifetime Healthcare, Lifetime Care, Med America Insurance Company, Rochester BlueCross BlueShield, and the St. Joseph's Neighborhood Center.

Robert A. Greene, M.D., is associate medical director of RIPA and a clinical assistant professor of medicine at the University of Rochester School of Medicine and Dentistry. He also maintains a small geriatrics practice. Dr. Greene received his undergraduate degree from Harvard University and his M.D. from the University of Pennsylvania. Dr. Greene has a special interest in physician-performance measurement. In addition, he is a member of the Symmetry (Ingenix) Medical Advisory Board, the Ambulatory Care Quality Alliance Performance Measurement Workgroup, and several technical expert panels on efficiency measurement. He has published widely on cost and quality-improvement issues. In 2005, Dr. Greene co-founded Focused Medical Analytics in order to bring the techniques developed at RIPA to a wider audience.

Editorial support was provided by Stephen J. Marcus.

EXECUTIVE SUMMARY

In their efforts to improve health care system efficiency, payers have become increasingly active in promoting strategies to improve the value of care. While initially focusing on quality, in order to address their continuing increases in costs of care, payers have recently added cost or efficiency measures to their programs. This paper reviews the four most popular strategies to address cost and efficiency: pay-for-performance, public reporting, consumer-directed health plans, and tiered or limited networks.

We begin by considering the strengths and weaknesses of the current methods of collecting and interpreting cost efficiency data, reporting results to interested parties, and providing incentives to shape provider behavior and patient outcomes. We then examine the four specific strategies in detail.

General Measurement Issues

Accuracy

The current trend in efficiency measurement is to calculate an index of physicians' costs relative to those of their peers. Efficiency indexes are built on the assumption that underlying claims data are correct, but there are a number of problems with that assumption. For example, on their visit billing forms, the doctors may use disease descriptions but not diagnostic codes. If the doctor writes, for example, "depression," a clerk must later decide whether this means minor depression, major depression, psychotic depression, reactive depression, or adjustment disorder with depressed mood, all of which are coded differently. In addition, because of the complex insurer-hospital contracts, which include payment formulas tied to volume and sites of service, the cost of any single hospitalization or procedure may be difficult or impossible to calculate. The result is estimated costs that may not accurately incorporate facility costs into the overall cost calculation. In episodes of care that involve hospitalization, those hospitalization costs are often a major determinant of total episode costs. The accuracy of each health plan's management of data is also crucial to reporting validity and should be transparent to the subjects of the reporting (i.e., providers) and those using the data to make decisions.

Attribution

There are many ways to assign a patient's costs to a practitioner, practice group, or integrated system, whether for population- or episode-based measures. In fact, different attribution rules can result in significantly different conclusions. Recent data suggest that some beneficiaries change assigned physicians from one year to the next, making assignment by individual physician inaccurate. Determining how patients' costs are

assigned may influence how some clinical decisions will be made and the degree to which physicians are likely to effectively collaborate. For example, if assignment of costs is based on the physician with the highest cost in the episode of care, the collaborating physicians will have incentive to have others order expensive tests. If, instead, all the physicians who contribute to more than 30 percent of the overall costs are assigned the cost of the episode, all the involved physicians will have incentive to talk with each other about how to minimize unnecessary care and costs.

Assignment of Cost

Practitioners reasonably argue that they can only be held accountable for the costs that they themselves engender. However, scoring on these direct or responsible costs creates an incentive to make *other* practitioners order the more expensive interventions, such as medications, imaging, consultations, or hospitalizations. This approach can cause delays in appropriate care because of disagreements over who should prescribe chronic medication refills or who should order an indicated MRI.

Incentive Program Goal

Stakeholders debate whether an efficiency measure should pay for achieving a target or for significant improvement. Those arguing for reaching targets say that to deliver true quality care, a particular level of care must be assured. Therefore, they conclude, one should be rewarded for reaching that target. On the other hand, if one employs an all-or-none model, those who were top performers before creating the incentive are rewarded, discouraging the lowest performers from investing in an improvement program with a target they are unlikely to reach, and encouraging providers to preferentially select patients already close to the target.

Actionability

For efficiency-reporting programs to be successful they must predictably pinpoint practitioners' overuse and misuse behaviors, those that add cost but do not improve desired outcomes. Simply providing a global rolled-up measure, such as a cost-efficiency index that compares them to a specialty average, is itself inefficient. For the busy practitioner, trying to find what to do differently—without the proper tools—results in wasted time, frustration, and, eventually, a lack of trust in the sponsoring organization. Meanwhile, in trying to become more efficient and improve their scores, they may reduce their treatment of necessary as well as unnecessary services. Finally, if the desired action is to move patients from less efficient to more efficient practitioners, program administrators should be certain there is excess capacity in the more-efficient group to accommodate the shift. If there is not capacity in the “more-efficient” group, patients will have to pay a

higher copay for care that is ranked as less-effective care. That situation can only lead to frustration and raise patients' concerns about their physicians' quality.

Pay-for-Performance

Pay-for-performance is designed to provide a direct link between payment and the performance of carefully selected services. The measures selected are intended to evaluate the practitioners' performance on certain combinations of cost/efficiency, quality, and a patient's experience of care. Although there is growing acceptance of pay-for-performance programs, its influence has been hampered by small incentives, lack of adequate post-intervention follow-up, and methodological problems with collecting and reporting the data.

We believe that insurers will come to understand that rather than unsuccessfully attempting to identify and reward the best practitioners, the most productive use of pay-for-performance programs is promoting behavioral change for a small set of carefully selected measures for which there is ample evidence of need for improvement. Organizing the professional community around a specific set of behaviors that are known to have a direct impact on outcomes creates incentive to improve. Although there is concern about incentives diverting attention from nonselected activities, if the measures chosen are of significant benefit, the time focused on them would be well spent.

Public Reporting

Public reports are designed to drive system change in three ways: using comparison data to motivate providers to improve performance, stimulating payers to reward quality and efficiency, and providing information to patients so that they can choose their care more wisely.

But physicians tend to see the public reporting of their individual quality and cost data as misleading and even threatening. Their concerns about accuracy, sufficient sample size, and the inherent judgment imposed by forced rankings—seemingly to motivate physicians to change through fear of humiliation or shame—put the intent of public reporting programs into question.

Data should of course be available to consumers to allow them to decide from whom to seek care and what outcomes they may expect. The challenge, however, is to present the data in a way that minimizes the judging of physicians and focuses on encouraging quality improvement. Prematurely providing misleading or insufficient information to consumers risks alienating practitioners while offering no real benefit to patients.

Consumer-Directed Health Plans (CDHP)

All variations of CDHP try to offer consumers broad choices of providers and services, coupled with greater information about—and liability for—their prices. In theory, if there is access to information on cost, care options, quality, and risks, patients will act as informed consumers in a competitive market, helping to manage their costs by taking more responsibility for their health care decisions.

But CDHP depends on patients' proper interpretation of reported data, an area in which even physicians have difficulty. Even more important, there is good evidence that patient cost-sharing can actually decrease the quality of care. As a patient's cost increases, fewer services, appropriate along with inappropriate, tend to be used.

Meanwhile, physicians are deeply concerned that they will spend valuable time looking up comparative fees, become more involved in negotiating rates, and have higher office expenses as bills are increasingly collected from patients' health savings accounts or out-of-pocket. To many practitioners, this model increases rather than decreases system inefficiencies. This is especially true now, when the data available to patients is simply inadequate to reliably improve decision-making.

Tiered and Limited Networks

Tiered networks utilize two mechanisms for controlling cost. The first is to encourage patients to switch from lower-value to higher-value physicians by offering lower copays for higher-value physicians or higher copays for lower-value physicians. The second mechanism is to encourage physicians to become more cost efficient in order to avoid the negative outcomes of being advertised as a lower-tier physician, with an attendant loss of patients.

Unfortunately, the criteria for selection in tiers can vary significantly. Selection is confounded, moreover, by issues related to geographic needs and sociodemographic variables, making the validity of these forced rankings suspect.

One obvious limitation of these approaches is access to providers. The concept of restricting access to a subset of providers necessarily implies that there is a pool of excess high-value providers from which to choose in that geographical area. In many communities, this is simply not the case, especially in rural and some inner-city areas.

Conclusions

The lack of demonstrated efficacy, the problems with trying to identify the precise qualities of a physician that lead to predictable increases in health care value, and the inherently judgmental nature of these four strategies suggest that public reporting, tiered and limited networks, and pay-for-performance programs that try to judge the overall quality of a physician are unlikely to fulfill their expectations. Results to date lead us to conclude that, in the health care industry, basing change on the foundation of win-lose relationships does not work. As we look to create a more effective and efficient health care system, it is critical that we not mistake programs to force behavioral change with successful quality-improvement processes.

In such a process, the physician's behavior is seen as contributing to a *system* of care. As that system is evaluated for how it delivers services, behaviors that help or hinder the attainment of desired outcomes may be identified. By focusing on specific units of behaviors rather than critiquing the overall performance of one individual physician, the focus remains on the need to improve rather than on placing blame.

We believe that the current focus on judgment programs has accomplished little and will continue to meet with limited success. These programs ignore some of the more important lessons that the business community has learned about true quality improvement. Successful improvement is rooted in partnership based on a commitment to achieve mutually agreed upon goals and a sense of shared benefit. Transparency must be delivered not only in the data reported to patients and purchasers but also in the processes that health plans use to make decisions about physician performance.

Successful transformation of our present system to one based on value requires the collaboration of multiple stakeholders. We focus on collaboration because effective partnerships are needed to effectively coordinate behaviors, provide a continuum of services, and ensure that evidence and patient-centeredness underlie medical decision-making. The current medical system's silos must be broken down and replaced with integrated programs that encourage and reward effective and efficient system solutions.

CURRENT APPROACHES TO IMPROVING THE VALUE OF CARE: A PHYSICIAN’S PERSPECTIVE

Dr. Sara Jerome is a 38-year-old family physician who lives in a community of 40,000 people. She works three days a week and carries a panel of 1,500 patients. She has two partners, each of whom works three days a week as well. The group employs Paul, a nurse practitioner who works five half-days a week. All four share after-hours call. Dr. Jerome admits patients to the university hospital 40 minutes away, where they are cared for by a hospitalist. Sicker patients are often transported there by ambulance. When her patients are discharged, about 70 percent of the time she receives information about final diagnoses, discharge medications, results of tests ordered, and suggestions for follow-up care.

The group employs five full-time office staff. John is the billing clerk/practice manager who has worked at the practice for 10 years. Yvonne is the receptionist, also on the job for 10 years, and Judy, George, and Becky are technicians who assist with patient care, call in refills, answer patient questions, and supply handouts requested by the practitioners.

Last week, Dr. Jerome received a “physician profile” from the dominant health plan in her community. A major element of the profile was an index of her relative cost-efficiency, adjusted for her case mix by episode-of-care methodology. She found the index hard to understand but thinks it reported that she cost 10 percent more than other plan doctors. An attached “drill-down” sheet suggested that Dr. Jerome’s higher expense resulted from her care of patients with diabetes, depression, heart failure, and esophagitis. Within these conditions, it appeared that facility costs were the dominant cost factor in her diabetes and heart-failure cases, medication prescribing was responsible for her costliness in esophagitis, and professional services were responsible for her increased costs in depression. Her patient-satisfaction and quality scores were above average, earning congratulations from the plan.

INTRODUCTION

In their efforts to improve system efficiency, payers have recently been taking cost measures, along with the traditional quality measures, into account—as in the above example. Physician behavior may be influenced both by financial and nonfinancial incentives.^{1,2,3} This paper will review the four most popular strategies currently being promoted: pay-for-performance, public reporting, consumer-directed health plans, and tiered or limited networks.

We begin by discussing the general strengths and weaknesses of collecting and interpreting data, presenting the results to interested parties, and providing incentives for provider behavior and outcomes. We then examine the strategies of pay-for-performance, public reporting, consumer-directed health care, and tiered or limited networks in some detail. We describe the logic behind each strategy, its results to date, its limitations, and its effects on practicing physicians. Our perspective is anchored in our experience at the Rochester Individual Practice Association (RIPA) from 1999 through 2006 in producing individual-physician performance reports and paying out (since 2002) an average of \$15 million yearly in a pay-for-performance program that operated at the individual-physician level. Our pay-for-performance program involved a 3,000-practitioner panel and was financed through our risk-withhold dollars (i.e., where a certain amount of each participating physician's fees is withheld and paid out based on the physician's performance on a set of predetermined quality, efficiency, and patient satisfaction measures) to which were added shared savings program dollars. The program was based on an HMO member base averaging 350,000 members per year in a nine-county service area. The size of the membership allowed us to provide profiling reports at the individual physician level.

The goal of each strategy is to increase the value of health care. However, the path has proven to be more difficult than imagined. Those from the business community hoping to transfer their business techniques to the medical world have been particularly frustrated.

Most current programs employ "judgment" as a principal focus, which means that data are being interpreted to create a summative—and subjective—evaluation of a physician's work. In Dr. Jerome's case, she was "judged" to be 10 percent more expensive than her colleagues, which presumably makes her a poorer-performing physician. This form of reporting makes a clear, though not necessarily accurate, inference about the physician's relative value.

At the other end of the performance-reporting spectrum is continuous quality improvement, in which the physician's behavior is seen as contributing to a system of care. As the system is evaluated for how it delivers services, behaviors that help or hinder the attainment of desired outcomes may be identified. By focusing on specific kinds of behaviors rather than the value of the individual based on a limited set of inputs, the focus remains on the system response to the need to improve rather than on defending one's value, career, or professionalism.

We believe that the current focus on judgment programs has accomplished little and will continue to meet with limited success, as these programs ignore some of the more

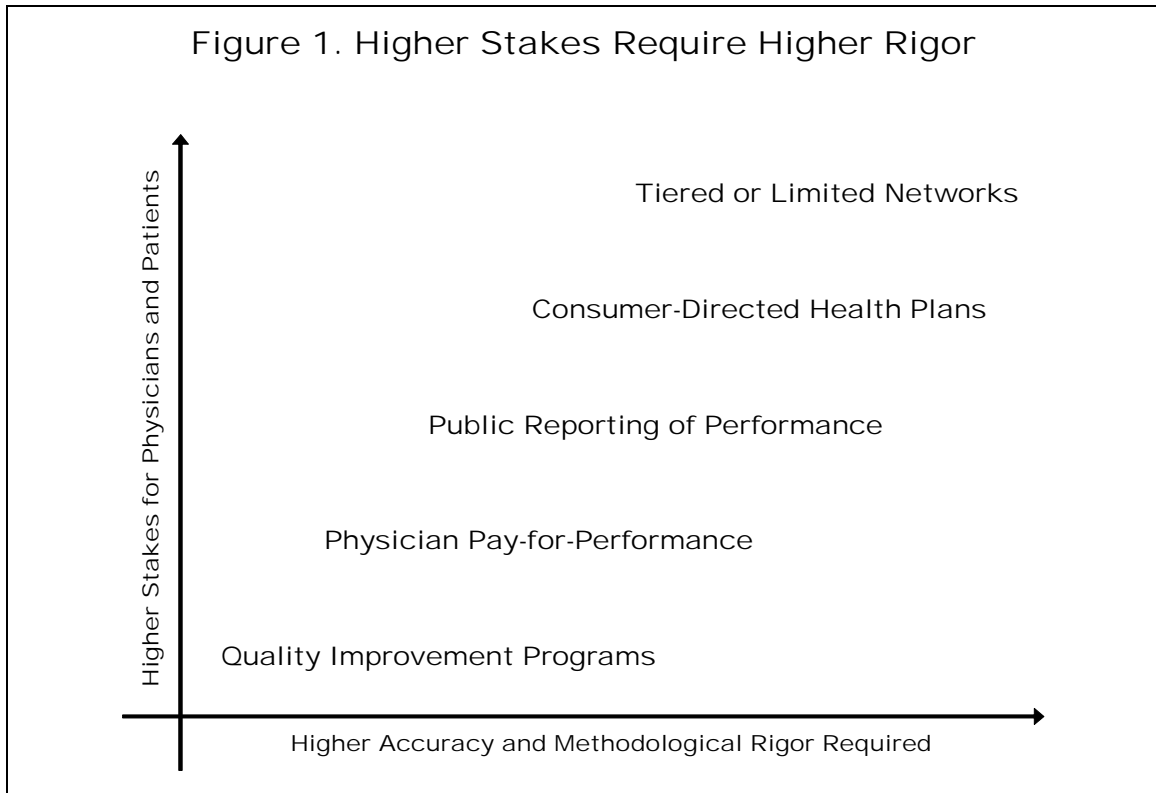
important lessons that the business community has learned about true quality improvement.⁴ In lieu of this focus, we encourage a value paradigm rooted in the core partnership values of honesty and respect. Transparency must be delivered not only in the data reported to patients and purchasers but also in the processes that health plans use to make decisions about physician performance.

Successful transformation of our present system to one based on value requires the collaboration of multiple stakeholders. We focus on collaboration because effective partnerships are needed to effectively coordinate behaviors, provide a continuum of services, and ensure that evidence and patient-centeredness underlie medical decision-making. The current medical system's silos must be broken down and replaced with integrated programs that encourage and reward effective and efficient system solutions.

EVALUATING CURRENT METHODS TO IMPROVE THE VALUE OF CARE

General Measurement Issues

Before incentive methods can be evaluated, the validity of the underlying performance-measurement system must be assessed. We believe that different validity and reliability standards apply to different incentive programs, depending on the size, type, and importance of the incentive. As the potential consequences of the program increase, the requirement for accuracy and validity increases as well.⁵ Figure 1 displays this relationship graphically. As we discuss each of four types of incentive systems, we will touch upon the consequences of insufficient accuracy and the perception of validity.



The IOM report *Performance Measurement: Accelerating Improvement* states that all such approaches “depend upon the availability of accurate, reliable, and valid performance measures.”⁶ In addition to these domains we would add attribution of the responsible physician(s), assignment of cost, focusing on reaching a target (as opposed to improvement), and actionability.

Accuracy

The current trend in efficiency measurement is to calculate an index of physicians’ costs relative to those of their peers. Based on such index results, high-stakes decisions—determining one physician’s copay relative to others, for example, or inclusion in a “high performance network”—may be made. These measures are referred to as efficiency indexes, performance factors, or observed-to-expected ratios.

Efficiency indexes are built on the assumption that underlying claims data are correct. But in Dr. Jerome’s practice, there are a number of problems with that assumption. First, on their visit billing forms, the doctors use disease descriptions but not diagnostic codes. For example, they write “depression.” It is John, the billing clerk, who determines the actual code number to put on the claim. Billers may do the best they can, but they usually do not have clinical training. In this case, John chooses 311.⁷ This turns out to be a very general code in which patients with minor depression, major depression,

psychotic depression, reactive depression, and adjustment disorder with depressed mood are included. Because the practice has never received feedback about its coding procedures and has no knowledge of how others code, Dr. Jerome and her staff believe their coding is appropriate. But, in fact, the distribution of costs of patients within the subclasses of depression easily could reflect length of treatment, the need for multiple medications, and concurrent psychiatric input. Therefore coding practices could make some cases look more or less expensive than average.

In addition, John is paid based on a percentage of charges received. Noting that current procedural terminology (CPT) 99214 visits pay \$84 while the less comprehensive CPT 99213 visits pay \$52, he encourages the group to spend a bit more time with each patient to maximize revenues (99213 corresponds to 15 minutes, 99214 to 25 minutes).⁸ The group has never received feedback on these types of billing practices, much less information on how important coding selection is to efficiency scoring.

The accuracy of each health plan's management of data is also crucial to reporting validity. But plan reporting of accuracy auditing is uncommon, if evaluated at all. It is becoming increasingly recognized that to be viewed as valid, the accuracy measurements should be transparent to the physicians being measured.⁹ To deal with errors when they occur, the reporting organization should outline how one can appeal the corresponding decisions, which should also help to prevent errors in the future. For example, until brought to the attention of the reporting team, there was considerable inaccuracy in the ordering-provider designation both for radiology and pharmacy in our community. Inaccurate data should limit the reporting agencies' confidence in using efficiency indexes for high-stakes decision-making regarding financial and nonfinancial incentives.

Attribution

There are many ways to assign a patient's costs to a practitioner, whether for population- or episode-based measures. Some include: a) the physician who generated the most costs for the patient's care for a particular diagnosis or episode; b) all the physicians who were each responsible for a minimum percentage (often 25%) of the episode's costs; c) all practitioners who have filed a claim for the patient; d) the practitioner who sees the patient most; and 5) the assigned primary care provider. These attribution models determine which costs are assigned to which doctors.

Interestingly, these attribution rules are rarely explained to the practitioners being evaluated. In Dr. Jerome's case, the office assigned the nurse-practitioner Paul to her, so all the care Paul delivers is billed under Dr. Jerome, whether she is supervising him in the office

that day or not. In addition, because the doctors all work part-time, they are continuously cross-covering each other's patients. As a result, two of the three doctors usually have less than 25 percent of the charges for an episode, so they are excluded from many of the episodes.

Different attribution rules yield significantly different results.¹⁰ Therefore it is important to know the attribution methodology used in a particular program and to understand that methodology's implications. The National Committee for Quality Assurance (NCQA) has proposed standards for cost attribution¹¹ and it requires disclosure of such methodology for its Quality Plus Program.¹²

Pham identified a further concern about the accuracy of attribution models.¹³ In examining continuity in Medicare recipients, she found that 33 percent of beneficiaries changed assigned physicians from one year to the next, and that only 35 percent of visits were to the beneficiary's assigned physician. Between 2000 and 2002, 46 percent of beneficiaries changed or were assigned to a new physician. Pham concluded that "the dispersion of patients' care among multiple physicians will limit the effectiveness of pay-for-performance initiatives that rely on a single retrospective method of assigning responsibility for patient care." These findings appear to challenge many efficiency determinations, as calculating an efficiency score usually requires a minimum two-year data-collection period to accumulate sufficient volume of care to create a valid practice assessment.

With the use of efficiency indexes—as part of incentive programs—in its infancy, it would seem prudent to determine the effects of several attribution models. Evaluators could then choose the one most likely to accomplish a program's goals before creating reimbursement or reporting models that significantly influence a practitioner's income or standing in the community.

Assignment of Cost

In a program based on episodes of care, efficiency indexes can be calculated from that portion of costs directly generated by the physician or from total episode costs. Direct costs include office visits to the practitioner, medications prescribed, tests ordered, or hospitalizations under his or her care. Total costs are those services generated by the practitioner in question plus all the costs generated by other providers during the episode. The decision to use directly generated or total costs has important consequences for practitioners' scores as well as ramifications for their uses.

Practitioners reasonably argue that they can only be held accountable for the costs that they themselves engender. However, scoring on direct costs creates an incentive to make other practitioners order the more expensive interventions. In such a system, the

practitioner can say, “I know how to appear less expensive—I’ll have *you* do the work.” This methodological choice has thus resulted in delays in appropriate care because of concerns over who should prescribe chronic medication refills or who should order the MRI, as that might worsen an individual’s efficiency score.

Our belief is that total cost is the most appropriate system measure. In this model, everyone has an incentive to work together most efficiently and effectively, as all are held jointly responsible for the costs of an episode of care.

On the other hand, one of the principles included in most guidelines for appropriate performance measurement is that the actions measured be within the control of the practitioner.^{14,15,16}

One solution to this contradiction might be to use total cost to drive “internal” evaluation, as opposed to external evaluation such as public reporting or tiering networks. Careful attention to the unintended consequences of each methodological decision, then, is a critical step in designing an effective program.

Focusing on Targets or Improvement When Measuring Cost-Efficiency

There has been considerable debate around whether an efficiency measure should focus on achieving a target or a percentage improvement. Those arguing for actually reaching targets say that to deliver true quality care, a particular level of care must be assured. Therefore, they conclude, one should be rewarded for reaching that target.

Hayward voices an important concern about aiming to reach the target threshold.¹⁷ He argues that if one employs an all-or-none model, the incremental effort to reach the target may result in little additional medical benefit while possibly introducing significant risk. For example, offering incentives for having all diabetics lower their HbA1c levels to less than 6.5 percent raises the risk of hypoglycemic episodes. Similarly, there may be unintended adverse effects from increasing drug dosages or ordering additional procedures to reach a target.

In another example, there is evidence of potential overuse of the medication epoetin by some dialysis programs to ensure patients maintain a hematocrit over 33.¹⁸ As in the previous case, neither the risks nor benefits of using epoetin specifically when hematocrit is close to 33 to reach the target are unknown.

The threshold model is carried even further by Health Partners in Minneapolis, which argues that the goal should be satisfying *all* selected quality measures for certain

clinical conditions.¹⁹ Health Partners terms this approach optimal care, and it credits providers' behaviors as such only when all targets for the clinical condition are reached. Others, including Donald Berwick of the Institute for Healthcare Improvement, have argued for the same approach.²⁰ Although it may work well in ensuring that the steps in an accepted process are carried out, the idea that all patients want what the medical establishment dictates seems quite paternalistic.

Moreover, providing a reward to a group that is already at or above a target accomplishes little. Recently, Young and colleagues found that the best predictor of meeting a target was having the highest pre-intervention scores.²¹ In our experience, the payer community justifiably argues against rewarding those who have not improved, given that the intent of the incentive is to encourage improved value and outcomes.

As a compromise, a hybrid model seems best for resolving most concerns. A system could reward providers for improvement toward the proximity of the threshold, using a continuous variable model. Medicare is proposing just such an approach for its hospital pay-for-performance program.²² The area in which the provider does the best, be it improvement or meeting the target, is reported and serves as the basis for payment.

In Dr. Jerome's case, she was judged a lower-performing doctor for diabetes care after it was determined that her rate of adherence to chronic-disease guidelines was 30 percent (the target was 65%). She decided that trying to reach a 65 percent target was not worth the effort. On the other hand, given a hybrid model, an increase of 50 percent would result in a reasonable return and entail an acceptable investment in time and money, in her opinion.

Actionability

It is 3:00 PM on a Saturday afternoon and Dr. Jerome is furiously reviewing records of patients whose care makes her cost-inefficient, according to the insurance plan's analysis of her efficiency index. For hypertension her pharmacy costs are high, while for respiratory infections her visit costs are high. She reviews the recent guidelines for hypertension and finds that she is adhering quite appropriately to the recommendations. That observation, together with what she has heard at the many CME programs she has attended, lead Dr. Jerome to believe that her medication selections are clinically sound.

As she looks at her notes for respiratory infections, she observes that when adults come in for colds, she often uses the visit to find out more about how they are doing generally, makes sure recommended screening is being performed, asks about smoking and alcohol use, and provides counseling where needed. This approach does generate longer visits, which apparently add to her lower cost-efficiency. “So,” she says, “I guess I’ll have folks come in separately for their preventive work.”

For efficiency-reporting programs to be successful, they must predictably pinpoint practitioners’ overuse and misuse behaviors. Simply providing a global rolled-up measure, such as a cost-efficiency index that compares them to a specialty average, is itself inefficient. For the busy practitioner, trying to find what to do differently—without the proper tools—results in wasted time, frustration, and, eventually, a lack of trust in the sponsoring organization.²³

If practitioners believe they are being evaluated on cost and the incentives are great enough, they will reduce service utilization in an effort to improve their scores. Unfortunately, in the process they may knowingly or unknowingly reduce both necessary as well as unnecessary services, thereby lowering quality. Little work has been done to determine how reporting of global scores influences appropriate or inappropriate behavior. However, Deming repeatedly noted that exhortations to improve do nothing, and that it is management’s job to find the specific actions that improve overall quality.²⁴

Rather than judge the individual’s overall performance, a more appropriate goal would be to determine the key drivers of unnecessary variations in care for common or costly conditions. Then a concerted effort could be made to ensure that necessary services are performed and unnecessary ones eliminated for each condition.

PROVIDING INCENTIVES FOR PHYSICIANS TO CHANGE

In this section, we review the four most common physician-incentive programs currently being used in the United States: pay-for-performance, public reporting, consumer-directed health plans, and tiering/limiting panels. To allow for comparisons between them, each discussion addresses the same set of issues: the logic behind the program, experience to date, limitations of the program, the practicing-physician perspective, and suggested future directions.

Pay-for-Performance

Logic Behind the Program

Of all the incentive programs currently deployed in the United States, pay-for-performance programs have been most widely used.²⁵ The reasoning behind pay-for-performance is

direct linkage of payment to the performance of designated services. The measures selected are intended to evaluate the practitioners' performance on certain combinations of cost, quality, and patient experience of care. Practitioners or medical groups then receive incentive payments based on the degree to which they either approach the measures' thresholds or improve their peer ranking for those measures.

Pay-for-performance programs expect that physicians will change their behavior in response to sufficient incentives; this model assumes that tying enough dollars to reported data will effect the behavioral change. Casalino surveyed general internists and found that they supported the concept of being remunerated for providing higher-quality care.²⁶ Seventy-five percent agreed that if quality measures were accurate, physicians could be evaluated and paid for their performance on those measures. That sentiment was recently affirmed through focus groups in Rochester, N.Y.²⁷ and in Massachusetts.

Experience to Date

Although there is growing acceptance of the appropriateness of payment based on effectiveness and efficiency of care, so far the evaluation of pay-for-performance programs' influence has been hampered by small incentives, lack of adequate post-intervention follow-up, and methodological problems with collecting and reporting the data. Higashi et al. have provided some evidence that the use of published targets for selected measures results in improved clinical outcomes.²⁸ Both Young and Rosenthal^{29,30} found minimal improvement in quality measures after initiation of pay-for-performance programs. More recently, Curtin et al. published the only study to date to evaluate return on investment (ROI) for a pay-for-performance program.³¹ That study demonstrated an ROI of 2.5 for the diabetes segment of a larger pay-for-performance program in upstate New York. Interestingly, the program demonstrated savings against the yearly trend in the plan's cost of diabetes care per member, even though the pay-for-performance measures focused on underuse. These results provide initial evidence that pay-for-performance programs can produce a positive ROI, as long as there is careful attention to measure selection, and that quality can be improved, with an attendant decreased cost.

Limitations of the Program

Recent studies have increasingly focused on some of the significant flaws of current pay-for-performance methodology. However, the consequences of such limitations are of even greater concern for the strategies of tiering, public reporting, and limiting panels, as will be discussed in later sections of this paper.

Hayward has raised four substantial concerns about pay-for-performance programs.³² First, he noted that patient comorbidities influence treatment targets, a concern echoed by Pogach.³³ An oversimplified outcome measure, be it quality or efficiency, in the absence of exclusion criteria and revised targets based on comorbidities may cause more harm than good. Second, the benefits of reaching targets are usually defined for a subset of patients with a given condition. Analogous to “indication creep” for medications and procedures,³⁴ there is a danger of extending the target to others with a less severe expression of the condition. Third, the marginal benefit of reaching a target, when the practitioner is already near the target, is unstudied; continued pursuit may even introduce potentially adverse effects. Fourth, patients able to approach the target—generally, those with less severe disease—become the more desirable patients; practitioners are essentially encouraged to recruit them, while avoiding those who are less likely to contribute to reaching the quality or efficiency target.

Finally, a number of organizations have advocated that physicians should be evaluated only on measures that are within their control.^{35,36,37} A number of the measures currently in use, including those focusing on cost-efficiency, may be significantly influenced by the population being treated or the health beliefs of those drawn to particular practitioners. The belief that adherence to a measure should be absolute suggests that the decision to proceed with a test or treatment is right or wrong. This dogmatic approach significantly limits a patient’s right to choose, a core component of patient-centeredness.

In summary, there is considerable concern that the overly simplistic approach currently being taken may mirror an all-too-common error in our medical system: implementing a costly and administratively complex program without sufficient evidence of its effectiveness. Considerable work is needed to more carefully examine the quality and efficiency measures that we impose on practitioners as well as the degree to which we ensure that the public is actually well served by our efforts.

Practicing-Physician Perspective

As noted earlier, physicians are becoming increasingly comfortable with the concept of pay-for-performance. On the other hand, there is evidence that they doubt the accuracy of current data and distrust the health plans’ motives in promoting the programs.³⁸ The degree to which physicians become effectively engaged in the process has much to do with the underlying values of the reporting group and the extent of the professional community’s participation in the development and maintenance of the program. Our experience suggests that physicians pass through predictable stages on the way to acceptance of pay-for-performance programs, similar to the stages in the dying process

described by Kubler-Ross.³⁹ Denial is the default coping skill, which is followed by anger, bargaining, and, potentially, acceptance. We have suggested that these in fact are the stages of coping with *any* significant change.⁴⁰

When groups reach the bargaining stage, available quality and efficiency measures do seem to stimulate changes in behavior. However, some have recently suggested that often the level of the individual physician is not the most effective point at which to intervene.⁴¹ Rather, the practice or system level is more appropriate and leads to greater success. Another predictor of success is the degree to which a practice is organized and able to hold employees accountable for completing assignments. That observation has been supported by the fact that the best predictor of reaching targets is being a high performer *before* the initiation of the program.⁴²

Physicians tell us that they want their individual data compared to that of peers. They want both quality and efficiency measures that are meaningful, and they especially want help in discovering how to improve their performance. In addition, they stress that the opportunity to improve should be available before their scores are reported to the public or used in judgments about the effectiveness and efficiency of their doctoring.

Future Directions

We believe that insurers will come to understand that the most productive use of pay-for-performance programs is in promoting behavioral change for a small set of carefully selected measures for which there is ample evidence of need for improvement. Depending on the sponsoring organization, the measures can be drawn from the domains of patient experience of care, quality, or cost-efficiency. Although there is concern about incentives diverting attention from nonselected activities, if the measures are chosen wisely they should be those of largest direct benefit to the community.

Critical to pay-for-performance as a quality-improvement program is the reporting system that gives feedback to practitioners. We agree with Trude, who argues that the most efficient and effective reporting system is one that measures and reports on community-wide data.⁴³ There are two reasons for endorsing this approach. First, a single reporting system increases the sample size, which makes it robust enough to offer meaningful and reliable data. Second, a current problem is that physicians receive conflicting reports from multiple plans, which confuses them about what they are supposed to be doing.⁴⁴

We also need considerably more research to define how to most effectively design a measurement system that captures the intricacies of patient comorbidities and balances preservation of individual patient autonomy with improving the population's health. In our view, it is premature to judge the overall quality of a physician based on the available information and the lack of evidence on judgments' accuracy or benefit. We predict that in the future, programs that encourage judgment—as opposed to improving community-wide performance—will be abandoned.

To create effective programs, the providers, purchasers, and insurers will have to collaborate more closely. They must establish win-win relationships in order to create programs whose incentives are aligned around truly improving the value of care for our entire population.

Public Reporting

The public reporting of health care costs and quality-performance data has long been viewed as a tool for improving the health care system's performance. As far back as 1984, CMS reported hospital mortality data publicly.⁴⁵ In 1990, New York State began reporting mortality rates from cardiac surgery.⁴⁶ After studying this program, Hibbard and Jewitt later wrote, "The current approach to health care reform attempts to harness the power of consumer choice. Great expectations are being placed on the consumers."⁴⁷

Logic Behind the Program

Public reports are designed to drive system change in three ways: using comparison data to motivate providers to improve performance, stimulating payers to reward quality and efficiency, and providing information to patients so that they can choose their care more wisely.

Experience to Date

Public reporting has met with varied success. Some of the variability is related to the kinds of data reported, their mode of presentation, and who they focus on. One of the early successes of public reporting was the New York report on cardiac surgery in the early 1990s,⁴⁸ followed by similar reports in Pennsylvania. These reports examined mortality data by hospital and cardiac surgeon. In New York, over three years there was a 41 percent decrease in coronary artery bypass graft mortality, thought to be largely the result of process improvement in hospitals.⁴⁹ A consequence of the program, foreshadowing concerns in future reporting programs, was the decision of some surgeons to avoid providing services to the highest-risk patients.⁵⁰

Although surgeons and their institutions have responded to public reporting, purchasers and consumers have not. It appears that the quality improvement derived from this program was driven by the hospital's or surgeon's desire to improve their public image.⁵¹ Influencing patients' decision-making appears to require as-yet-unidentified factors.

In another example, *Quality Counts*—a 2001 report on hospital performance by The Alliance, a large employer purchasing cooperative in Madison, Wis.—compared performance measures for hospitals receiving no report, hospitals that received their report privately, and those whose report was disseminated publicly. Further analysis revealed that while both publicly and privately reported hospitals improved significantly compared to hospitals not receiving reports, the difference in average performance changes between publicly reported and privately reported institutions was not statistically significant.⁵²

As in the cardiac surgery experience, the study also found that public reporting had little effect on market share; only 4 percent of those exposed to the report used it to choose or recommend a hospital.

The Minnesota Community Measurement Project and the Massachusetts Health Quality Partners are two other groups actively reporting publicly on performance at the medical-group level. At this time, there are no published reports on market-share changes attributable to these efforts. But based on its experience with public reporting, Minnesota's Institute for Clinical Systems Improvement (ICSI) has made three observations.⁵³ First, public reports supplement organizations' efforts if the reports also address methods for quality improvement. Second, some historical evidence demonstrates that improved process and outcomes may be tied to public reporting. Third, there is no evidence that consumers and purchasers use performance data to make choices when they purchase health care.

CMS is currently starting to promote increased reporting of performance data.⁵⁴ At present it is focused on long-term care and home health. In 2007, it is launching voluntary physician reporting, with added reimbursement earmarked for those who report clinical data. All reported data will be made available to the public.

Limitations

Many of the limitations of pay-for-performance apply to public reporting as well. A major one has been insufficient data to make statistically valid conclusions about any given provider. With small sample sizes per provider, it is frequently impossible to determine whose performance is different from the average. And because public reporting puts

physicians' reputations on the line, the professional and emotional stakes are higher than with pay-for-performance programs. Providers are likely to increasingly game the system to improve score or ranking or to prevent embarrassment. For example, they may avoid sicker patients, have others involved in a patient's care order tests or prescribe medicine, or alter visit or procedure coding to shift patients from a more scrutinized condition to a less scrutinized one.

The public reporting of performance indexes (cost or quality relative to average) is especially problematic. Because most physicians practice in smaller groups with limited IT and system design resources, the sharing of best practices in medicine is particularly important. But indexes are forced rankings that create an inherent competition, which discourages colleagues from actually sharing best practices.

Practicing-Physician Perspective

As with pay-for-performance, providers' response to public reports has largely followed the predictable emotional response to change.⁵⁵ The Alliance in Wisconsin, in the initial report on its program, clearly described anger and disbelief on the part of hospital staff.⁵⁶ In a survey of general internists by Casalino et al. regarding public reporting and pay-for-performance,⁵⁷ the majority of respondents were supportive of financial incentives but only 45 percent were supportive of group-level reporting, and support decreased to 32 percent for individual performance.

Maleyeff⁵⁸ has postulated that one reason for professionals' resistance to public reporting is Deming's principle of avoiding "management by numbers."⁵⁹ Deming speaks to individuals' reward or punishment based on data over which they have little or no direct control. This is consistent with ICSI's observation that individual reports need to be connected to actionable improvement.

Massachusetts Health Quality Partners (MHQP) has collaborated with the Massachusetts Medical Society (MMS) to better understand how to report data successfully. They sponsored eight focus groups with local physicians,⁶⁰ which showed once again that providers were clearly concerned about being evaluated on items outside their control. Examples were vaccination refusal based on patient preference and the inability to afford a medication copay.

There was also a strong belief that performance data should be shared with those physicians being evaluated before the information is distributed to the public. This would give them a chance to first improve their scores or outcomes. It was also noted that if the goal

of public reporting truly is improvement, any deficiency noted should be accompanied by actionable recommendations on how the individual's performance might improve. Last, practitioners expressed great concern about the unintended consequences mentioned earlier regarding cardiac care. If publicly released evaluations are negatively affected by sicker or disadvantaged patients, the incentive is to decrease these populations' access to care.

Future Directions

It is clear that physicians see the public reporting of their individual quality and cost data as very threatening. And their concerns about accuracy, sufficient sample size, and the inherent judgment imposed by forced rankings put the intent of public reporting programs into question. When added to the fact that patients don't change their behavior based on the released data, the only reason for public reporting appears to be motivating physicians to change through fear of humiliation or shame.

Data should of course be available to consumers to allow them to decide from whom to seek care and what outcomes they may expect. The challenge, however, is to present the data in a way that minimizes the judging of physicians and focuses on encouraging quality improvement. Attention to the design and language of public reporting programs, and the inclusion of physicians in their development and rollout, will determine the efficacy of public reporting on the quality and cost of care.

Consumer-Directed Health Plans

Consumer-directed health plans (CDHP) are a recent insurance innovation aimed at addressing the rising costs of health coverage. The term "consumer-directed health plan" is somewhat confusing, as it really refers to many different benefit designs. The central concept in all these plans, however, is public reporting combined with a high-deductible health plan (HDHP) and tax-advantaged medical spending accounts.

Logic Behind the Program

All variations of CDHP try to offer consumers broad choices of providers and services, coupled with greater information about—and liability for—their prices. In theory, if there is access to information on cost, care options, quality, and risks, patients will act as informed consumers in a competitive market, helping to manage their costs by taking more responsibility for their health care decisions. This model has evolved in direct response to concern over the negative cost implications of "moral hazard," the tendency for those not exposed to the cost of care to increase demand and overuse services.⁶¹ CDHP shifts decision-making about the use of services from physicians or plans directly to the patient.

At present, a tax-advantaged account is usually either a health reimbursement account (HRA) or a health savings account (HSA). HRAs are funded and owned by the employer, while HSAs are owned by the employee and are portable between employers. HSAs and HRAs differ also in restrictions on funding, qualification, and rollovers, which are beyond the scope of this discussion.

The anticipated outcome of the consumer-directed design is that market forces will spur innovation, resulting in the delivery of higher-quality and more efficient health care. After all, in market economies it is consumer demand that typically drives innovation and results in lower prices and improved products. Economic theory holds that even when only a relatively small portion of all purchasers actively “shops” for value, the market pressure exerted on providers is significant.

Experience to Date

HSAs, formally established by the Medicare Modernization Act of 2003, have actually been part of the tax code since 2001. But as of 2006, only 7 percent of companies offering health benefits even offered an HDHP, with or without a savings account.⁶² Deductibles for these plans are five times greater than those of PPOs and 20 times more than HMOs. If a subscriber’s deductible exceeds the balance of the savings account, the subscriber may be liable for the difference.

The subscriber’s potential costs seem to be a significant barrier to informed-consumer acceptance of CDHP. This behavior is predicted by Daniel Kahneman’s “prospect theory,” which postulates that potential losses are perceived to have greater economic impact than potential gains of the same actual monetary value.⁶³ Thus it is no surprise that employees chose an HSA consumer-directed plan only 19 percent of the time when other options are offered. The majority of employees with an HSA-eligible plan (53%) choose it because they have not been offered other plan options. Woolhandler and Himmelstein agree, arguing that CDHP may be good for the healthy and wealthy, but otherwise it is not to the consumer’s objective advantage to join such a plan.⁶⁴

At this time there is only preliminary data about CDHP. The employer’s total premium cost is reported to be less for these plans than for traditional plans, but when all medical expenses are calculated, including the savings-account expenses, total costs are essentially equal.⁶⁵ This would characterize cost-shifting rather than cost-reduction. Unpublished reports from Mercer, Aetna, and Humana have reported, however, that the cost increases in these plans are from one-third to one-half of the average trend in no-

HDHP products. Beyond cost to the employer, overall system value cannot be evaluated for these plans, given their low market penetration and short duration of use.

Limitations of CDHP

The limitations of CDHP again include those of pay-for-performance and public reporting. Further, CDHP depends on patients' proper interpretation of reported data, an area in which even physicians have difficulty.

Even more important, there is good evidence that patient cost-sharing can actually decrease the quality of care.⁶⁶ The RAND Health Insurance Experiment (HIE) of the 1970s and 1980s looked at the impact of increasing cost-sharing on care provided. Not surprisingly, as patient cost increased, fewer services were used. The most troubling aspect of that work was that reductions were equal both for appropriate and inappropriate services. Although services avoided by the total study population had little aggregate impact on health status, reduced utilization in the subgroup of sick and poor patients resulted in inferior outcomes.⁶⁷

A recent study of emergency department (ED) use by patients with HDHPs showed similar results.⁶⁸ Of relevance to this discussion, the decline in ED use for CDHP occurred only after an initial visit. It would appear that after the sticker shock of receiving the bill, members found alternatives. ED use declined 24 percent overall, with a much greater decline in visits categorized as low-severity. However, as in the HIE, patients in the lower-income categories experienced a 25 percent reduction in visits for high-severity conditions. The implication is that those with fewer financial means in plans with greater out-of-pocket exposure begin to behave like the uninsured—that is, they seek less medical care than insured individuals do, and have poorer short-term health outcomes.

The differential response to high deductibles—a function of economic status—is particularly relevant when looking at CDHP enrollment. Current enrollees have higher incomes and appear to be in somewhat better health than those in traditional plans. This raises the issue of selection bias, which confounds any studies of the cost benefits of CDHP. Moreover, a greater adverse effect on lower-income patients raises the additional issue that adoption of CDHP violates the Institute of Medicine principle that health care should be socially equitable.

Practicing-Physician Perspective

Evaluation of physician views on CDHP design has been quite limited. A recent study by Pham, which examined how often physicians talked with patients about cost, found that

while discussions on drug cost were quite common, discussions related to testing and care setting were infrequent.⁶⁹ Communication skills retraining will be needed both for office staff and practitioners in order to incorporate patients' questions, in already time-constrained office visits, about choices of care.

A more fundamental concern about CDHP has been raised by Berenson.⁷⁰ He suggests that CDHP would undermine the tradition of trust between doctor and patient and replace it with a vendor model based on "caveat emptor."

Shifting the fulcrum of the physician–patient relationship to appropriateness of resource utilization could dramatically change the use of time in the office visit and, at least initially, create role confusion. In the face of data suggesting that needed services will be deferred, physicians are deeply concerned. More specifically, they fear that they will spend valuable time looking up comparative fees, become more involved in negotiating rates, and have higher office expenses as bills are increasingly collected from patient's health savings accounts or out-of-pocket. To many practitioners, this model increases rather than decreases system inefficiencies.

Future Directions

CDHP attempts to impose a structural response to concerns that, in the current health care system, no one is motivated to evaluate the true worth of an intervention or to eliminate those that do not add value. Speaking in economic terms, the cost of an intervention has to be weighed against the value obtained if those funds were spent elsewhere. If patients bear responsibility for the actual costs of services, one can argue that "moral hazard" will be avoided as they look at the costs and benefits more critically.⁷¹

The degree to which this model decreases costs and influences utilization has yet to be determined. No one has yet assessed the tradeoff between the dollar savings and patient harm caused by deferring appropriate care. But for the model to be successful, much work will need to be done in training practitioners to integrate cost-effectiveness discussions into medical visits and other conversations with patients. A negative impact on the processes of care would be an important unintended consequence, which practitioners must be vigilant in avoiding.

CDHP relies on market forces to change the system. To the extent that those forces are not applicable to health care, CHDP will falter. A market is successful at driving value when there are product or service choices, where demand is predictable, and where the consumer has the ability to compare. The current system is lacking in all these areas.

Thus for CDHP to really succeed in improving the value of health care, at a minimum there needs to be valid and readily accessible information on cost, outcome, and risk. We simply do not have that data to give patients at present. As noted above, lack of information to make decisions is already a point of enrollee dissatisfaction. Even if data were available, it is not known whether patients would be ready to incorporate the implications in order to improve care. Tu and May looked at consumer behavior in existing self-pay markets where information was readily available and there was relatively little time pressure on office visits.⁷² Despite this best-case scenario, they found that patient choices were based largely on referral recommendations rather than objective measures.

The continued concerns about affordability and quality will ensure continued discussion on how to engage patients in their own care. Value-based insurance design has begun to be suggested as an alternative to relying solely on patients to make decisions. This model suggests that an individual's financial liability should not be as large as it is in CDHP but based instead on medical evidence that evaluates the relative value of each intervention. High-value treatments would have little or no cost to the individual, while low-value interventions would have higher costs.⁷³

Tiered and Limited Networks

Logic Behind the Program

As consumers rejected managed care in the mid-1990s, payers and health plans scurried to identify other methods to control their spiraling costs.⁷⁴ Tiered networks, which were designed in response, utilize two mechanisms for controlling cost. The first is to encourage patients to switch from high-cost to lower-cost physicians by offering lower copays for lower-cost physicians or higher copays for higher-cost physicians. The second mechanism is to encourage physicians to become more cost-effective in order to avoid the negative outcomes of being advertised as a lower-tier physician, with an attendant loss of patients.^{75,76}

Experience to Date

Drapier and colleagues have recently offered a market scan of the early experiences with tiering and high-performance networks.⁷⁷ They report that the most common model uses tiered-provider levels with corresponding enrollee cost-sharing differentials (varied copays). In a minority of programs, the information is provided to members for information only. Most often, plans have targeted specialists for these networks and programs. With little evidence of efficacy, the criteria for selection in tiers can be quite varied. Selection is confounded by issues related to geographic needs and sociodemographic variables, making the validity of the forced rankings suspect.

Some plans mix quality and cost measures, while some focus more on cost alone. As in the pay-for-performance world, the lack of clarity of purpose may well lead to problems. Quality measures that address underuse would increase short-term costs, blunting or nullifying the cost-saving effect of efficiency measures. Since there is little correlation demonstrating performance on quality and cost measures, as the tiering criteria attempt to be more global, the cost saving benefits may be diluted by the addition of quality measures into the equations.

However accomplished, the politics of physician selection is a significant intervention in the delivery of care. But again, there are no published data on the outcomes of these programs. Marketing materials by insurers and consultants encouraging such programs suggest savings; however, it seems ironic that programs justifying their existence with calls for transparency of cost data have not subjected their methodology and results to the very transparency they espouse.

Limitations

One obvious limitation of these approaches is related to access to providers. The concept of limiting access to a subset of providers necessarily implies that there is a pool of excess providers from which to choose in that geographical area. This is simply not the case in many parts of the country, especially in rural and some inner-city areas. Recent evaluations document many areas where physician supply, even without limitations, is marginal at best.

The impact of limited access is reflected in problems and compromises of programs already implemented. Physicians who normally would be excluded are “invited” onto panels that are based not on qualification but on geographic access. Also, once a few employers in a community effectively shift patients from “less” to “more” effective/efficient providers, those latter providers may fill and need to close their practices to all others. Thus the rest of the community is left paying a higher copay to see physicians designated as weaker performers. The result is likely to be a large group of unsatisfied employees or consumers.

Another problem with this approach is that physicians typically practice in groups. But many, including the present authors, have seen as much variation within groups as between groups. Difficulties are created, both for patients and physicians, regarding whether networks are tiered by group or individual practitioner. If the tiers are created at the group level, there will be a regression to the mean and limited potential for results. If

individual physicians are placed in different tiers, patients and practices are hampered by disruptions in established patterns of coverage and care.

For example, if Dr. A is a low-copay physician but is away, and a patient used to seeing her needs to be seen, the patient may visit with Dr. B, who has a high (\$25) copay rather than the customary top-tier copay of \$10. Worse, imagine a physician assistant who is supervised by physicians in different tiers. If the patient sees the physician assistant on, say, Dr. B's day rather than Dr. A's day, does he or she pay the higher copay for seeing the same person?

From the technical perspective, there is considerable concern that measures used to define effective/efficient physicians are neither valid nor reliable enough to make such important judgments.⁷⁸ Current measures are quite limited and are complicated by several factors: a lack of data to demonstrate the benefits of making treatment decisions that may cause only minor changes in scores,⁷⁹ the absence of a broad set of measures to examine the depth and breadth of a physician's care, and the relative absence of measures to assess the quality of a practitioner's patient-centeredness.

For example, even if a physician's diabetes and cost scores are excellent, how does one also factor in his or her availability for urgent needs, willingness to see patients outside of office hours, and capacity to respond to the grief and sorrow of coping with significant illness? As we focus on "hard" measures, we have yet to determine what effect that has on the other aspects of the care process.

Even if all such measures were perfected, physician performance falls on a bell-shaped curve. There are relatively few outstanding physicians and equally few poor ones. Trying to influence the behavior of those in the middle risks tampering with the system by moving patients based on chance events. Deming argues that such interventions actually increase variation in performance.⁸⁰

Practicing-Physician Perspective

So far, tiering has been met with significant physician dissatisfaction. In St. Louis and Washington State, physician organizations sued insurers, who later rescinded their programs.^{81,82} In Massachusetts, physicians have protested tiering methodologies through the Massachusetts Medical Society. Concerns are focused on a number of issues highlighted earlier, including inaccurate data, unreliable cost measures, the absence of meaningful risk adjustment, and the use of measures on which the physician has minimal influence.⁸³ The lack of data justifying the intervention, combined with the highest stakes

of any of the programs discussed in this paper, has thus led to significant pushback by the physician community.

Physicians' major criticisms of tiering programs are both technical and interactional. Provided with any detailed information behind their scores, physicians quickly recognize the technical limitations described above. Tiered systems inherently violate principles of relationship-centered administration, which have been shown to be a key in physician engagement and quality improvement.^{84,85}

The other aspect of the physician community's response is an interpersonal one. These views are summed up by the comments of four Massachusetts physicians, obtained during the previously mentioned focus groups sponsored by MHQP and MMS and conducted by Dr. Beckman. These comments are published with permission of both organizations.⁸⁶

Physician 1: I don't think anyone minds the game when the game is supposedly to improve the quality of care. If there is a game to improve the quality of care, sign me up, I'll play that game. But when the game is unfair, and the rules are all askew, that is where the problem is. If you can fix it to make the data real, bring it on.

Physician 2: Whenever I purchase something, I do a lot of research. Either *Consumer Reports* or Amazon because the information is reliable, it's reproducible and from an independent body, not the manufacturer. The problem with the quality measures with insurers, the measures are good but the person doing it sometimes has an incentive that sometimes conflicts and that's the concern we have as physicians. It's not credible and it doesn't take into account the other variables.

Physician 3: Good care, we say, is inexpensive care. If the patient does well and gets out of the hospital quickly, that's more inexpensive than someone who does poorly. The system does deserve to have its money used efficiently. I don't think we are as efficient in this country as we can be with the utilization of our resources and I guess that's what it's all about. But it certainly shouldn't be punitive and we should all be involved in the way it's decided, and maybe that's why we are here tonight.

Physician 4: They (the physician hospital organization) are now taking the step to meet with specialists and actually show them data on the rates of their surgical outcomes. We see numbers. When you see yourself at the bottom of the list, you try and work your way up.

Generally, the focus-group physicians were disheartened and angered by what appeared to them to be an undeserved attack on individual physicians who had little advance notice of the impending assault, and, more important, no opportunity to improve.

Such feelings were mirrored as well in Drapier and colleagues' summary of early experiences.⁸⁷ Their conversations with practitioners identified the following concerns about the tiering model: lack of communication from the plans about the program; being uninformed about their designations within the high-performance system; plans' lack of explanation about how performance was assessed; unavailability of data that would help drive improvement; the accuracy and reliability of available data; insufficient sampling size on which to base these important decisions; and the lack of standardization in methodologies (a practitioner might be in different tiers for different plans). United's experience in St. Louis and Regence's experience in Washington State have been the two most visible examples of physician pushback against these programs.^{88,89}

Future Directions

The future for tiering and limited panels is not at all clear at the present. As noted, this approach is not useful in most small and medium-sized markets, where there is no pool of excess physician capacity; and it is methodologically problematic in all markets, whatever their size. Some may argue, however, that even when used for small numbers of practitioners, the impact will be surprisingly great because of practitioners' fear of being publicly chastised or excluded from a panel.

To date, these programs have not provided physicians with the information needed to improve their performance. As a result, it is difficult to imagine how significant savings will be achieved. And a potentially greater problem is that, in the absence of meaningful information, physicians will reduce both necessary and unnecessary services as part of a more generic cost-cutting response.

Tiered networks combine the highest stakes with the most difficult technical issues, and at a time when the measures available to make the determinations of value are inadequate for the task. Adding the effects of such a highly judgmental model on physicians' sense of professionalism and work satisfaction, we believe that the future of tiered networks is quite limited.

CONCLUSIONS

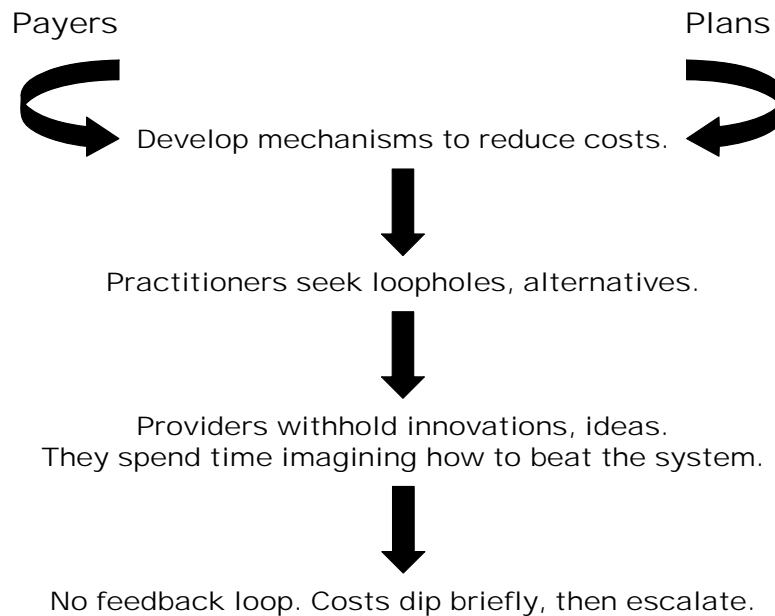
As long as the costs of care continue to escalate, payers will keep searching for strategies to successfully control their health care cost trend. Their urgent need to decrease costs draws them to simple, far-reaching programs such as pay-for-performance, public reporting,

consumer-directed health plans, and tiered networks. All such programs have some intuitive attractiveness. Pay-for-performance reflects the idea that “you get what you pay for.” And it seems logical to direct patients to the more efficient effective physicians through public reporting, the application of consumer incentives, or simply eliminating lower-scoring practitioners from a network.

Unfortunately, in the field of medicine, as in other complex human endeavors, the path to improved quality is strewn with simple logical concepts that failed. The lack of demonstrated efficacy, the problems with trying to identify the precise qualities of a physician that lead to predictable increases in health care value, and the inherently judgmental nature of all these interventions suggest that they are unlikely to fulfill their expectations. For example, the judgmental nature of tiered and limited networks, together with the high stakes of such interventions, suggests especially strong physician resistance.

An examination of these approaches’ less-than-glowing results leads us to the conclusion that, in the health care industry, basing change on the foundation of win-lose relationships does not work (Figure 2). To successfully transform the health care system for producing greater value, we need to organize it by agreeing on fundamental principles and focusing on a clear set of goals. There is no question that as a nation we have to find a successful strategy to improve the value of care we provide. But so far, the challenge has been to align the key stakeholders rather than simply transfer dollars from one silo to another.

Figure 2. The Win-Lose Cycle



Source: H. B. Beckman, Rochester Individual Practice Association, Mar. 2004.

To succeed, the perspectives of providers, consumers, payers, and insurers have to be elicited and then molded into programs that respect all participants. Moreover, each group must be held accountable for its performance in achieving the desired health outcomes and financial results.

It is important, as we look to create a more effective and efficient system, that we not mistake programs to force behavioral change with a successful quality-improvement process. What is more, attempts to externally impose motivation through the use of incentives have had very mixed results; indeed, they have spawned their own set of problematic behaviors.⁹⁰ The key to success is to have motivational methodology that works synergistically with a continuous quality-improvement process.

In our view, tiering and high-performance networks, and even poorly designed public reporting programs, fail to meet these criteria. These programs require the identification of winners and losers, but *all* practitioners are essential to maintaining access to the system. Creating a population of physicians who feel humiliated and disrespected, yet still have to help support the system, is a recipe for disaster. We are beginning to see that outcome playing out in a number of communities already.

Consumer-directed health plans have theoretical strengths, but they are years away from creating sufficient data and delivering it in an understandable format. The reality, for the time being, is that we will be asking medically naïve consumers to use inadequate data to drive decisions that are difficult even for their more experienced and knowledgeable physicians. Pretending that the reason for promoting these programs is informed choice is simply irresponsible. CDHP is being promoted to shift costs, pure and simple.

Pay-for-performance has the potential to be part of the solution, for two reasons: improvement can be built into the performance program, and robust data systems can help practitioners identify specific behaviors that may be causing their inefficiency or lower quality. The risks of employing pay-for-performance programs lie in the degree to which they default to judgment programs and attempt to determine who are the “best” and “worst” physicians. Rather, pay-for-performance should carefully determine the key foci that require change in a given community at a given time. Those foci should become the measures selected, and results should be reported to the public, payers, and providers. To the degree that pay-for-performance programs promote improvement, offer actionable feedback, and encourage the reduction of overuse, underuse, and misuse, they will improve the care process.

In the final analysis, Dr. Jerome, like almost all of her colleagues, wants to be reasonably compensated for providing high-quality and cost-effective care. She will feel engaged and respected if she is offered accurate and valid data about her patient population, clear guidance about what to change in order to be considered a better practitioner, and specific action items regarding how to improve. Given such a win-win relationship and effective leaders, the medical community will be an active and enthusiastic participant in transforming our health care system.

RECOMMENDATIONS

General Measurement Issues

Accuracy

- Health plans should report internal accuracy-audit results to evaluated practitioners.
- Reporting organizations should outline how one can appeal decisions based on inaccurate data and prevent future errors.
- Informational programs should be developed for practitioners so that they understand the methodology of efficiency indexing and can ensure accurate data inputs.

Attribution

- Evaluated practitioners should be informed of the attribution methodology used by plans to determine efficiency scores.
- For Medicare patients, individual physician-efficiency scoring should be abandoned until concerns regarding continuity issues are resolved.
- Determine the effects of several attribution models and then choose the one most likely to accomplish a program's goals. This should be done before creating reimbursement or reporting models that significantly influence practitioners' income or community standing.

Assignment of cost

- Evaluated practitioners should be informed of the cost methodology employed by plans to determine efficiency scores.
- To promote collaboration and a sense of combined responsibility for patients' care, total episode costs—as opposed to direct costs—should generally be used for determining a physician's or physician group's efficiency.

Focusing on targets versus improvement

- Evaluated practitioners should be informed if a target, improvement, or a hybrid of both is the criterion.
- A hybrid is preferred, and the proposed Medicare hospital methodology should be used as a model.

Actionability

- Determine the key drivers of unnecessary variation in care for common or costly conditions within a specialty. Then make the reduction of that variation a goal of the reporting program.
- Incentivize practitioners to reduce unnecessary variation by specifically identifying overuse behaviors and modes of behavioral change.

Providing Incentives for Physicians to Change

Pay-for-performance (Note: Recommendations do not imply endorsement of these programs by the authors or The Commonwealth Fund.)

- Pay-for-performance programs to motivate behavioral change should be based on a small set of carefully selected measures. They should be employed only when there is ample evidence of, and agreement among all stakeholders about, the need for improvement.
- Develop methods to address patient severity.
- Provide compensation based on improvement as well as on reaching a target. Employ continuous variables for payment so as to discourage incremental treatment in pursuit of minimal clinical improvement.
- Ensure that measures address the correct organizational level. Report at the individual-physician level when the desired outcome is within the control of the physician and the measure is clinically meaningful.
- Measure performance based on community-wide data.

Public reporting

- Report only accurate and audit-proven data to the public. It is not appropriate to use vendors' black-box technology without audited results.
- Data should be reported to physicians well in advance of reporting to the public. In that way, inaccuracies and flawed scoring methodologies can be identified and corrected before physicians are publicly judged.
- Reported measures should be actionable at the provider level, or at whatever level is most appropriate. Efficiency indexes and other roll-up models are too inaccurate to use for high-stakes reporting.

Consumer-directed health plans

- At least initially, do not market CDHP to low-income families or to patients with complex medical problems.
- Before expanding CDHP to consumers/employees, improve the quality of data and the transmission of information.
- Before purchasing CDHP, members should undergo communication-skills training and be educated about the use of data and their role in this new model of decision-making.
- Given already time-constrained office visits, practitioners and staff will need to be trained to address patients' questions about choices of care.

Tiered and limited networks

- Exempt physicians in underserved areas and underpopulated specialties from the tiering process.
- Develop pilot programs to ensure that there are sufficient accuracy of selection and actual returns to warrant the effect on professionals' lives and careers.
- Quality and patient-experience-of-care measures should be included in the data set that determines tier and network membership.
- Tiering programs should only be implemented after physicians have been apprised of the measures to be used and given a reasonable period of time to improve their performance.
- Tiers should be based on absolute performance—not on forced rankings—so that all who achieve the desired level of performance can receive credit.

NOTES

- ¹ C. Schoen, S. K. H. How, I. Weinbaum, J. E. Craig, Jr., and K. Davis, *Public Views on Shaping the Future of the U.S. Health System* (New York: The Commonwealth Fund, Aug, 2006).
- ² L. G. Sandy and D. M. Bazarko, “Advancing Quality in a Consumer-Directed World: ‘May You Live in Interesting Times,’” *American Journal of Managed Care*, May 2005 11(5):287–88.
- ³ Institute of Medicine, *Rewarding Provider Performance: Aligning Incentives in Medicine* (Washington, D.C.: National Academies Press, 2007).
- ⁴ W. E. Deming, *Out of the Crisis* (Cambridge: Massachusetts Institute of Technology Press, 1982).
- ⁵ R. A. Greene, H. B. Beckman, G. H. Partridge et al., *Review of the Massachusetts Group Insurance Commission Physician Profiling and Network Tiering Plan: A Report to the Massachusetts Medical Society* (Waltham, Mass.: Massachusetts Medical Society, 2006). Available at http://www.massmed.org/AM/Template.cfm?Section=Pay_for_Performance&TEMPLATE=/CM/ContentDisplay.cfm&CONTENTID=16760. Accessed June 11, 2007.
- ⁶ Institute of Medicine, *Performance Measurement: Accelerating Improvement* (Washington, D.C.: National Academies Press, 2006).
- ⁷ American Medical Association, *International Classification of Diseases, ICD-9-CM-2007, Physician, Vols. 1 and 2, 9th revision* (Chicago: AMA, 2006).
- ⁸ American Medical Association, *Current Procedural Terminology, CPT 2007, Professional Edition* (Chicago: AMA, 2006).
- ⁹ National Committee for Quality Assurance, *NCQA Quality Plus Program for Managed Care Organizations and Preferred Provider Organizations* (Washington, D.C.: NCQA, 2005).
- ¹⁰ Medicare Payment Advisory Commission, *Report to the Congress: Increasing the Value of Medicare. Chapter 1: Using Episode Groupers to Assess Physician Resource Use* (Washington, D.C.: MedPAC, June 2006). Available at http://www.medpac.gov/publications/congressional_reports/Jun06_Ch01.pdf. Accessed July 3, 2007.
- ¹¹ National Committee for Quality Assurance, “HEDIS 2007 Technical Specifications for Physician Measurement (Expanded Scope) Draft for Public Comment” (Washington, D.C.: NCQA, 2007). <http://www.ncqa.org/Programs/HEDIS/PublicComment/main.htm>. Accessed February 9, 2007. This site was only accessible during the comment period and the posting then removed.
- ¹² MedPAC, *Report to the Congress, Chapter 1*, 2006.
- ¹³ H. H. Pham, D. Schrag, A. S. O’Malley et al., “Care Patterns in Medicare and Their Implications for Pay for Performance,” *New England Journal of Medicine*, Mar. 15, 2007 356(11):1130–39.
- ¹⁴ Massachusetts Medical Society, *Guidelines for Measuring, Reporting, and Rewarding Physician Performance* (Waltham, Mass.: Massachusetts Medical Society, May 2005).
- ¹⁵ American Medical Association, *Pay for Performance: A Physician’s Guide to Evaluating Incentive Plans* (Chicago: AMA, 2005).
- ¹⁶ E. S. Fisher, “Paying for Performance—Risks and Recommendations,” *New England Journal of Medicine*, Nov. 2, 2006 355(18):1845–47.
- ¹⁷ R. A. Hayward, “All-Or-Nothing Treatment Targets Make Bad Performance Measures,” *American Journal of Managed Care*, Mar. 2007 13(3):126–28.

- ¹⁸ M. Thamer, Y. Zhang, J. Kaufman et al., “Dialysis Facility Ownership and Epoetin Dosing in Patients Receiving Hemodialysis,” *Journal of the American Medical Association*, Apr. 18, 2007 297(15):1667–74.
- ¹⁹ “HealthPartners Awards \$1.4 Million in Physician Bonuses,” news release (Bloomington, Minn.: HealthPartners, Nov. 9, 2005). Available at <http://www.healthpartners.com/portal/3059.html>.
- ²⁰ T. Nolan and D. M. Berwick, “All-or-None Measurement Raises the Bar on Performance,” *Journal of the American Medical Association*, Mar. 8, 2006 295(10):1168–70.
- ²¹ Communication from Gary Young, J.D., Ph.D., Professor, Boston University School of Public Health, received May 2007.
- ²² CMS Hospital Value-Based Purchasing Workgroup, “Options Paper: Second Public Listening Session” (Washington, D.C.: CMS, 2007). Available at <http://www.cms.hhs.gov/AcuteInpatientPPS/downloads/HospitalVBPOptions.pdf>.
- ²³ H. Beckman, A. L. Suchman, K. Curtin et al., “Physician Reactions to Quantitative Individual Performance Reports,” *American Journal of Medical Quality*, May-June 2006 21(3):192–99.
- ²⁴ W. E. Deming, *Out of the Crisis*, 1982.
- ²⁵ G. Baker and B. Carter, *The Evolution of Pay-for-Performance Models for Rewarding Providers* (Washington D.C.: Atlantic Information Services, 2004).
- ²⁶ L. P. Casalino, G. C. Alexander, L. Jin et al., “General Internists’ Views on Pay-for-Performance and Public Reporting of Quality Scores: A National Survey,” *Health Affairs*, Mar./Apr. 2007 26(2):492–99.
- ²⁷ Beckman, Suchman, Curtin et al., “Physician Reactions,” 2006.
- ²⁸ T. Higashi, P. G. Shekelle, J. L. Adams et al., “Quality of Care Is Associated with Survival in Vulnerable Older Patients,” *Annals of Internal Medicine*, Aug. 16, 2005 143(4):274–81.
- ²⁹ M. B. Rosenthal, R. G. Frank, Z. Li et al., “[Early Experience with Pay-for-Performance](#),” *Journal of the American Medical Association*, Oct. 12, 2005 294(14):1788–93.
- ³⁰ G. J. Young, M. Meterko, H. Beckman et al., “Effects of Paying Physicians Based on Their Relative Performance for Quality,” *Journal of General Internal Medicine*, June 2007 22(6):872–76.
- ³¹ K. Curtin, H. Beckman, G. Pankow et al., “Return on Investment in Pay for Performance: A Diabetes Case Study,” *Journal of Healthcare Management*, Nov.-Dec. 2006 51(6):365–74.
- ³² Hayward, “All-Or-Nothing Treatment Targets,” 2007.
- ³³ L. M. Pogach, A. Tiwari, M. Maney et al., “Should Mitigating Comorbidities Be Considered in Assessing Healthcare Plan Performance in Achieving Optimal Glycemic Control?” *American Journal of Managed Care*, Mar. 2007 13(3):133–40.
- ³⁴ R. A. Harrington and E. M. Ohman, “The Enigma of Drug-Eluting Stents: Hope, Hype, Humility, and Advancing Patient Care,” *Journal of the American Medical Association*, May 9, 2007 297(18):2028–30.
- ³⁵ Massachusetts Medical Society, *Guidelines for Measuring, Reporting, and Rewarding*, 2005.
- ³⁶ AMA, *Pay for Performance: Physician’s Guide*, 2005.
- ³⁷ Fisher, “Paying for Performance,” 2006.
- ³⁸ Beckman, Suchman, Curtin et al., “Physician Reactions,” 2006.
- ³⁹ E. Kubler-Ross, *On Death and Dying* (New York: Scribner, 1969).

- ⁴⁰ Beckman, Suchman, Curtin et al., “Physician Reactions,” 2006.
- ⁴¹ Greene, Beckman, Partridge et al., *Review of Massachusetts Physician Profiling and Network Tiering Plan*, 2006.
- ⁴² Communication from Gary Young, J.D., Ph.D., Professor, Boston University School of Public Health, received May 2007.
- ⁴³ S. Trude, M. Au, and J. B. Christianson, “Health Plan Pay-for-Performance Strategies,” *American Journal of Managed Care*, Sept. 2006 12(9):537–42.
- ⁴⁴ Beckman, Suchman, Curtin et al., “Physician Reactions,” 2006.
- ⁴⁵ J. M. Colmers, *Public Reporting and Transparency* (New York: The Commonwealth Fund, Jan. 2007).
- ⁴⁶ R. Steinbrook, “Public Report Cards—Cardiac Surgery and Beyond,” *New England Journal of Medicine*, Nov. 2, 2006 355(18):1847–49.
- ⁴⁷ J. H. Hibbard and J. J. Jewett, “Will Quality Report Cards Help Consumers?” *Health Affairs*, May/June 1997 16(3):218–28.
- ⁴⁸ Steinbrook, “Public Report Cards,” 2006.
- ⁴⁹ M. N. Marshall, P. G. Shekelle, S. Leatherman et al., “The Public Release of Performance Data: What Do We Hope to Gain? A Review of the Evidence,” *Journal of the American Medical Association*, Apr. 12, 2000 283(14):1866–74.
- ⁵⁰ J. H. Burack, P. Impellizzeri, P. Homel et al., “Public Reporting of Surgical Mortality: A Survey of New York State Cardiothoracic Surgeons,” *Annals of Thoracic Surgery*, Oct. 1999 68(4):1195–2000.
- ⁵¹ J. H. Hibbard, J. Stockard, and M. Tusler, “Does Publicizing Hospital Performance Stimulate Quality Improvement Efforts?” *Health Affairs*, Mar./Apr. 2003 22(2):84–94.
- ⁵² J. H. Hibbard, J. Stockard, and M. Tusler, “Hospital Performance Reports: Impact on Quality, Market Share, and Reputation,” *Health Affairs*, July/Aug. 2005 24(4):1150–60.
- ⁵³ G. Mosser and S. Scheitel, “Public Reporting of Health Care Performance in Minnesota: A Position Paper of the Board of the Institute for Clinical Systems Improvement” (Bloomington, Minn.: Institute for Clinical Systems Improvement, Nov. 2004).
- ⁵⁴ Centers for Medicare and Medicaid Services. <http://www.cms.hhs.gov/default.asp> Accessed July 3, 2007.
- ⁵⁵ Beckman, Suchman, Curtin et al., “Physician Reactions,” 2006.
- ⁵⁶ Hibbard, Stockard, Tusler, “Does Publicizing Hospital Performance,” 2003.
- ⁵⁷ Casalino, Alexander, Jin et al., “General Internists’ Views,” 2007.
- ⁵⁸ J. Maleyeff, F. C. Kaminsky, A. Jubinville et al., “A Guide to Using Performance Measurement Systems for Continuous Improvement,” *Journal for Healthcare Quality*, July–Aug. 2001 23(4):33–37.
- ⁵⁹ Deming, *Out of the Crisis*, 1982.
- ⁶⁰ Communication from Barbra Rabson, Massachusetts Health Quality Partners, and Elaine Kirshenbaum, Massachusetts Medical Society, received October 19, 2007.
- ⁶¹ M. Gladwell, “The Moral-Hazard Myth: The Bad Idea Behind Our Failed Health-Care System,” *The New Yorker*, Aug. 29, 2005, pp. 44–49.

⁶² Henry J. Kaiser Family Foundation and Health Research Educational Trust, *Employer Health Benefits 2006 Annual Survey* (Washington, D.C.: KFF/HRET, 2006). Available at <http://www.kff.org/insurance/7527/index.cfm>. Accessed July 3, 2007.

⁶³ D. Kahneman, P. Slovic, and A. Tversky, *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge, U.K.: Cambridge University Press, 1982).

⁶⁴ S. Woolhandler and D. U. Himmelstein, "Consumer Directed Healthcare: Except for the Healthy and Wealthy It's Unwise," *Journal of General Internal Medicine*, June 2007 22(6):879–81.

⁶⁵ KFF and HRET, *Employer Health Benefits 2006 Annual Survey*, 2006.

⁶⁶ K. Davis, [*Will Consumer-Directed Health Care Improve System Performance?*](#) (New York: The Commonwealth Fund, Aug. 2004).

⁶⁷ J. P. Newhouse, "Consumer-Directed Health Plans and the RAND Health Insurance Experiment," *Health Affairs*, Nov./Dec. 2004 23(6):107–13.

⁶⁸ J. F. Wharam, B. E. Landon, A. A. Galbraith et al., "Emergency Department Use and Subsequent Hospitalizations Among Members of a High-Deductible Health Plan," *Journal of the American Medical Association*, Mar. 14, 2007 297(10):1093–1102.

⁶⁹ Pham, Schrag, O'Malley et al., "Care Patterns in Medicare," 2007.

⁷⁰ R. A. Berenson, "Which Way for Competition? None of the Above," *Health Affairs*, Nov./Dec. 2005 24(6):1536–42.

⁷¹ Gladwell, "Moral-Hazard Myth," 2005.

⁷² H. T. Tu and J. H. May, "Self-Pay Markets in Health Care: Consumer Nirvana or Caveat Emptor?" *Health Affairs* Web Exclusive (Feb. 6, 2007):w217–w226.

⁷³ R. S. Braithwaite and A. B. Rosen, "Linking Cost Sharing to Value: An Unrivaled Yet Unrealized Public Health Opportunity," *Annals of Internal Medicine*, Apr. 17, 2007 146(8):602–05.

⁷⁴ M. Edlin, "Tiers Keep Costs in Tow," *Managed Healthcare Executive*, May 1, 2004, pp. 33–36.

⁷⁵ Greene, Beckman, Partridge et al., *Review of Massachusetts Physician Profiling and Network Tiering Plan*, 2006.

⁷⁶ D. M. Berwick, B. James, and M. J. Coye, "Connections Between Quality Measurement and Improvement," *Medical Care*, Jan. 2003 41(1 Suppl.):I-30–I-38.

⁷⁷ D. A. Drapier, A. Liebhaber, and P. B. Ginsburg, *High-Performance Health Plan Networks: Early Experiences*, Issue Brief No. 111 (Washington D.C.: Center for Studying Health System Change, May 2007).

⁷⁸ Greene, Beckman, Partridge et al., *Review of Massachusetts Physician Profiling and Network Tiering Plan*, 2006.

⁷⁹ Hayward, "All-Or-Nothing Treatment Targets," 2007.

⁸⁰ Deming, *Out of the Crisis*, 1982.

⁸¹ R. Kazel, "Tiered Physician Network Pits Organized Medicine vs. United," *American Medical News*, Mar. 7, 2005. Available at <http://www.ama-assn.org/amednews/2005/03/07/bil10307.htm>. Accessed June 18, 2007.

⁸² "Regence Shelves Narrow Network in Response to Physician Concerns," *Managed Care Week*, June 5, 2006.

⁸³ Greene, Beckman, Partridge et al., *Review of Massachusetts Physician Profiling and Network Tiering Plan*, 2006.

⁸⁴ M. C. Beach and T. Inui, "Relationship-Centered Care: A Constructive Reframing," *Journal of General Internal Medicine*, Jan. 2006 21(1 Suppl.):S3-S8.

⁸⁵ D. G. Safran, W. Miller, and H. Beckman, "Organizational Dimensions of Relationship-Centered Care: Theory, Evidence, and Practice," *Journal of General Internal Medicine*, Jan. 2006 21(1 Suppl.):S9-S15.

⁸⁶ Communication from Barbra Rabson, Massachusetts Health Quality Partners, and Elaine Kirshenbaum, Massachusetts Medical Society, received October 19, 2007.

⁸⁷ Drapier, Liebhaber, Ginsburg, *High-Performance Health Plan Networks*, 2007.

⁸⁸ Kazel, "Tiered Physician Network," 2005.

⁸⁹ "Regence Shelves Narrow Network," 2006.

⁹⁰ E. L. Deci and R. Flaste, *Why We Do the Things We Do: Understanding Self-Motivation* (New York: Grosset/Putnam, 1995).

RELATED PUBLICATIONS

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

[*Creating Payment Systems to Accelerate Value-Driven Health Care: Issues and Options for Policy Reform*](#) (September 2007). Harold D. Miller.

[*Value-Driven Health Care Purchasing: Four States that Are Ahead of the Curve*](#) (August 2007). Sharon Silow-Carroll and Tanya Alteras.

[*Lessons from Local Access Initiatives: Contributions and Challenges*](#) (August 2007). Karen Minyard, Deborah Chollet, Laurie Felland, Lindsey Lonergan, Chris Parker, Tina Anderson-Smith, Claudia Lacson, and Jaclyn Wong.

[*An Analysis of Leading Congressional Health Care Bills, 2005–2007: Part II, Quality and Efficiency*](#) (July 2007). Karen Davis, Sara R. Collins, and Jennifer L. Kriss.

[*Quality Matters: Payment Reform*](#) (July 2007). Newsletter.

[*Aiming Higher: Results from a State Scorecard on Health System Performance*](#) (June 2007). Joel C. Cantor, Cathy Schoen, Dina Belloff, Sabrina K. H. How, and Douglas McCarthy.

[*Pay-for-Performance in State Medicaid Programs: A Survey of State Medicaid Directors and Programs*](#) (April 2007). Kathryn Kuhmerker and Thomas Hartman.

[*Creating Accountable Care Organizations: The Extended Hospital Medical Staff*](#) (December 5, 2006). Elliott S. Fisher, Douglas O. Staiger, Julie P. W. Bynum, and Daniel J. Gottlieb. *Health Affairs* Web Exclusive (*In the Literature* summary).

[*The 2nd Annual EBRI/Commonwealth Fund Consumerism in Health Care Survey, 2006: Early Experience With High-Deductible and Consumer-Driven Health Plans*](#) (December 2006). Paul Fronstin and Sara R. Collins.

[*Why Not the Best? Results from a National Scorecard on U.S. Health System Performance*](#) (September 2006). The Commonwealth Fund Commission on a High Performance Health System.

[*Value-Based Purchasing: A Review of the Literature*](#) (May 2003). Vittorio Maio, Neil I. Goldfarb, Chureen T. Carter, and David B. Nash.

[*How Does Quality Enter into Health Care Purchasing Decisions?*](#) (May 2003). Neil I. Goldfarb, Vittorio Maio, Chureen T. Carter, Laura Pizzi, and David B. Nash.