

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12424
INCENTIVES FOR PHYSICIANS

Introduction

Apple Pickers or Federal Judges: Strong versus Weak Incentives in Physician Payment

There is broad agreement that paying for value in health care makes more sense than simply paying for volume (VanLare and Conway 2012; Porter and Lee 2015). Reflecting this agreement, in January 2015, Department of Health and Human Services Secretary Sylvia Mathews Burwell announced that HHS has set a goal of tying 85 percent of all traditional Medicare payments to quality or value by 2016 and 90 percent by 2018, and of having 30 percent of all traditional Medicare payments be made through alternative payment models, such as accountable care organizations (ACOs) or bundled payment arrangements, by the end of 2016, and 50 percent by the end of 2018 (Burwell 2015). The goals of shifting payments to alternative payment models and linking remaining fee-for-service payments to quality are also strongly reflected in the MACRA legislation enacted in 2015 (Steinbrook 2015).

However, there are significant challenges in ensuring that this shift will lead to improved patient outcomes (Werner and Dudley 2012; Ryan and Press 2014). In a lecture given 20 years ago, Northwestern University Economics Professor Burton Weisbrod suggested that consideration of the differences between payment systems for apple pickers and for federal judges may help inform consideration of how best to pay physicians. Payment for apple pickers is the archetypal example of a strong incentive system—apple pickers are paid for each apple they pick. In contrast, payment for federal judges is the archetype of a weak incentive system—federal judges have lifetime tenure and their pay is not related to any measure of performance. It makes sense to use a strong incentive system for apple pickers because it is relatively easy to measure whether they are delivering the desired product. The cost of measurement is low, and the apple picker has few opportunities to game the

measurement system. The product of apple picking is well defined and precise, so there is little concern that rewarding apple picking diverts attention away from other productive behaviors that employers would want apple pickers to carry out. In contrast, we want federal judges to produce high-quality justice efficiently but are not confident that we have good ways of measuring whether they do so. We are concerned that attempts to measure and reward efficiency—for example, measuring the time to complete a case—might result in judges skimping on the quality of justice they dispense. And we do not currently have the capacity to reliably measure judicial quality. As a result, instead of strong incentive systems for judges, we rely on the selection process and the norms of judicial professionalism to produce the desired outcomes.

Health—the product that we want produced in an efficient, high-quality manner by hospitals, physicians, and other health care providers—has many features that make it more like federal judging than like apple picking. For example, we want physicians to use their special knowledge and experience to exercise good judgment in choosing which actions to perform and to do this professionally in the best interests of their patients. Much of the work that physicians do is complex and highly dependent on the specific needs of particular patients, making it difficult to construct useful measures aggregated across patients. In addition, attempts to measure and reward some areas of performance run the risk of diverting physicians' attention from other areas that may be more important but are more difficult to measure. This concern is stated starkly in the terms of principal-agent theory in a classic textbook:

If an employee is expected (by the principal who pays the employee) to devote time and effort to some activity for which performance cannot be measured at all, then incentive pay cannot be effectively used for any other activities (e.g., easy-to-measure ones) that the employee controls. (Milgrom and Roberts 1992)

Measuring physician performance—for example, the rates at which a physician's patients appropriately have mammograms (a “process” measure) or the rates at which a physician's patients have potentially preventable ambulatory care-sensitive admissions (an “outcome measure”)—can provide actionable information for physicians and the organizations in which they

Address correspondence to Richard Kronick, Ph.D., Agency for Healthcare Research & Quality, 540 Gaither Road, Rockville, MD 20815; e-mail: richard.kronick@ahrq.hhs.gov. Lawrence P. Casalino, M.D., Ph.D., is with the Department of Healthcare Policy and Research, Weill Cornell Medical College, New York, NY. Andrew B. Bindman, M.D., is with the Departments of Medicine and Epidemiology & Biostatistics and the PRL-Institute for Health Policy Studies, University of California, San Francisco, San Francisco, CA.

work. But when a reward (financial or reputational, through public reporting) is attached to the measurement, it gives physicians an incentive to focus their attention on that measure as well as an incentive to avoid patients (e.g., the sick, the poor, and the poorly educated) who may lower their scores on that measure. If the incentive is relatively small, then physician professionalism may be sufficient to avoid these negative outcomes. However, there is concern that providing financial incentives for good performance may actually decrease physician professionalism (Wynia 2009). It is an open empirical question whether it is possible to “right-size” incentives for quality, making them large enough to induce improved performance on the output being measured, without making them so large as to crowd out attention from all of the many other things that we want physicians to do which are not being measured or rewarded.

Furthermore, it is difficult or statistically impossible to accurately measure the performance of individual physicians in most specialties with regard to key health outcomes—such as ambulatory care-sensitive admissions—because the number of cases may be too small (Nyweide et al. 2009), and because of difficulties adjusting for the characteristics (e.g., health status, poverty) of the physician’s patients that may affect the outcomes (Lipstein and Dunagan 2014). To varying degrees the papers in this volume distinguish between providing incentives for physicians and providing incentives for the organizations, such as medical groups, hospitals, and ACOs, in which physicians work. The quality of health care is determined not only by the decisions and actions of individual physicians but also by the investments in creating infrastructure and systematic processes to improve quality made by these organizations. Organizations that make these investments should at least have the opportunity to recoup the funds they have invested. Financial incentives for good performance can provide this opportunity.

Although there are reasons to be concerned about the results of rewarding physicians for measured performance, there are equally strong reasons to be concerned about a payment system that relies solely on professionalism to produce desired outcomes. Fee-for-service payment results in strong incentives to provide additional services, regardless of the effect of the additional services on outcomes that patients value. However, despite these incentives, much care of proven value is not delivered, care is often not responsive to patient preferences, and care is not as safe as it could be or should be. In the context of fee-for-service, the lack of a link between payment and value not only creates incentives for the delivery of care of unproven value, but also does not create incentives for the delivery of care that patients value highly.

Capitation and bundled payment systems reduce or eliminate incentives to deliver care of unproven value but do not, by themselves, create incentives for the delivery of care of proven value, and they create the additional concern that physicians will have incentives to stint on needed services and avoid patients most in need of care. Capitated and bundled payment systems therefore create a very strong imperative to measure and reward value.

As we move in the direction of paying for value, there are a number of key questions that need to be addressed:

1. What aspects of provider performance should be measured?
2. What types of incentives should be used?
3. How strong should the incentives be?
4. When should strong incentives be used and when should they not be used?
5. What can be done to encourage improved performance in areas of care where strong incentives may not be appropriate?
6. How do the answers to these questions differ when the underlying payment system is fee-for-service, capitation, bundled payment, or a shared-savings approach?
7. How do the answers to these questions vary when the organization being paid is a large medical group as opposed to a solo physician or a small group of practitioners?
8. How do the answers to these questions vary when the question is asked about how an organization (e.g., a large medical group or hospital) provides incentives for its physicians—rather than how a payer (e.g., Medicare or a health insurer) provides incentives for physicians?

To address these questions and to provide a framework for future research, the Agency for Healthcare Research and Quality (AHRQ) commissioned the five papers featured in this special issue of *Health Services Research*. The papers focus on incentives for physicians. The HSR editors have graciously permitted the papers to be longer than the usual journal article, making it possible for the authors to explore their topics in depth. The papers are not intended to be comprehensive systematic reviews; rather, they are intended to provide the best thinking of experienced researchers and to provide both practical policy suggestions as well as suggestions for key research that is needed. The papers focus, respectively, on suggesting (1) a conceptual framework for the use of financial incentives in health care; (2) key implications of the evidence to date on pay for performance and public reporting in health care; (3) innovative uses of measures of patient experience; (4) alterna-

tives to the use of strong incentives to foster improvements in provider performance; and (5) specifying a research agenda for the use of incentives in health care.

In the first paper, Conrad explains principal-agent theory and uses this theory, and to a lesser extent findings from behavioral economics, to thoroughly explore possible ways to use incentives in health care (Conrad 2015). He argues that empirical findings about incentives in health care, though still limited in quality and scope, are consistent with the predictions of principal-agent theory and behavioral economics. Conrad is particularly concerned with when and how payers such as health insurers, Medicare, and Medicaid should provide incentives to the organizations (e.g., medical groups and hospitals) in which physicians work and when and how they should be provided to individual physicians. This concern is particularly relevant as Medicare prepares to introduce performance-based payment for individual physicians.

In the second paper, Roland and Dudley provide a thoughtful, selective review of what is known about pay for performance and public reporting of physician performance in medical care, focusing particularly on experience in the United States and in the British National Health Service (Roland and Dudley 2015). They discuss the advantages and disadvantages of alternative payment systems, including fee-for-service, salary, and capitation, and argue that the unintended, undesirable consequences of each of these systems can be mitigated by the use of appropriate pay for performance and public reporting incentives. They argue that the effects of P4P and public reporting programs depend importantly on the basic payment system (fee-for-service, salary, capitation) to which they are added. Roland and Dudley conclude by suggesting principles that policy makers could use to maximize the benefits and minimize the unintended consequences of using incentives for physician performance.

In the third paper, Schlesinger, Grob, and Shaller (2015) identify four forms of “patient-reported information”: (1) patient-reported outcomes measuring self-assessed physical and mental well-being, (2) patient experience surveys, (3) narrative accounts describing encounters with clinicians in patients’ own words, and (4) complaints/grievances signaling patients’ distress when treatment or outcomes fall short of expectations. They argue that these forms differ in crucial ways, and that each should be distinctively measured, deployed, and linked with financial incentives. They are concerned that if public policies are not attentive to patients’ perspectives, stronger financial incentives for clinicians are likely to threaten aspects of care that patients most value. The paper by Schlesinger et al. is particularly notable for its detailed, novel discussion of the possible uses of patient narratives as means to improve

the value of care, and for the authors' suggestions on ways to increase the value and use of such narratives.

In the fourth paper, Berenson and Rice draw from insights in behavioral economics to argue that at present financial incentives for physicians are over-emphasized and that this may compromise physicians' "intrinsic motivation" as professionals (Berenson and Rice 2015). They suggest placing more reliance on tools that appeal more to physicians' intrinsic motivation. When an important quality problem does not lend itself to accurate measurement, it should be addressed with strategies using technical assistance and collaboration among providers. They make a variety of specific suggestions for such strategies, including promoting quality improvement collaboratives among providers and providing greater recognition for local quality improvement projects, providing feedback to providers from other providers as well as from patients, giving providers confidential feedback on their performance compared with that of other providers, strengthening Conditions of Participation for Medicare and Medicaid, and targeting prior authorization requirements to physicians who are outliers in potentially inappropriate use of expensive and/or dangerous procedures and imaging studies rather than to all physicians.

In the fifth paper, Luft (2015) provides specific suggestions for policy-relevant research on physician incentives, focusing on areas where little is known and where one or two research findings may have important implications for policy makers. The paper focuses on research that can be accomplished within 3–5 years within the range of research budgets provided by federal research and private foundations. Luft argues that there is no one-size-fits-all solution to providing physicians with incentives, and he suggests research that would help determine the situations in which certain incentives are better than others. Luft argues that fee-for-service payment is unlikely to disappear any time soon, and he suggests research to improve the way fee-for-service payments are used. He also focuses on research relevant to episode-based incentives, a subject of considerable relevance as CMS moves toward mandatory episode-based payment for some conditions.

Following the five papers, Kao (2015) interprets the papers from the perspective of the American Medical Association's efforts to understand physician work motivation, focusing particularly on the possible effects of incentives on the balance between extrinsic and intrinsic motivation in physicians. In the final commentary, Glied (2015) makes two arguments. First, she suggests that measures that are useful for accountability to external organizations such as CMS are likely to be quite different from measures that are useful for internal quality improvement and management efforts. Second, she

encourages policy makers to agree on a clear understanding of how and why quality improvement occurs in health care, noting that the main dynamic leading to quality improvement in most areas of the economy—that is, creative destruction—has not been much evident in health care.

The papers in this volume provide a thorough discussion of the challenges in using strong financial incentives as a mechanism to reward the production of high-quality health care, suggestions about mechanisms to deal with these challenges, and suggestions about research needed to make further progress. The question is not whether to use incentives, but how best to do so. AHRQ is committed to supporting research that will help public and private payers to more effectively pay for value.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: This project is funded by the Agency for HealthCare Research and Quality.

Disclosures: None.

Disclaimers: None.

Richard Kronick
Lawrence P. Casalino
Andrew B. Bindman

REFERENCES

- Berenson, R. A., and T. Rice. 2015. “Beyond Measurement and Reward: Methods of Motivating Quality Improvement and Accountability.” *Health Services Research* 50 (6 Pt 2): 2155–86.
- Burwell, S. M. 2015. “Setting Value-Based Payment Goals—HHS Efforts to Improve U.S. Health Care.” *New England Journal of Medicine* 215 (372): 897–9.
- Conrad, D. A. 2015. “The Theory of Value-Based Payment Incentives and Their Application to Health Care.” *Health Services Research* 50 (6 Pt 2): 2057–89.
- Glied, S. 2015. “Strong versus Weak Incentives: The Role of Policy, Management, and Theory in a New Research Agenda.” *Health Services Research* 50 (6 Pt 2): 2223–8.
- Kao, A. C. 2015. “Driven to Care: Aligning External Motivators with Intrinsic Motivation.” *Health Services Research* 50 (6 Pt 2): 2216–28.
- Lipstein, S. H., and W. C. Dunagan. 2014. “The Risks of Not Adjusting Performance Measures for Sociodemographic Factors.” *Annals of Internal Medicine* 161 (8): 594–6.

- Luft, H. S. 2015. "Policy-Oriented Research on Improved Physician Incentives for Higher Value Health Care." *Health Services Research* 50 (6 Pt 2): 2187–215.
- Milgrom, P., and J. Roberts. 1992. *Economics, Organization, and Management*. Englewood Cliffs, NJ: Prentice-Hall.
- Nyweide, D. J., W. B. Weeks, D. J. Gottlieb, L. P. Casalino, and E. S. Fisher. 2009. "Relationship of Primary Care Physicians' Patient Caseload with Measurement of Quality and Cost Performance." *Journal of the American Medical Association* 302 (22): 2444–50.
- Porter, M. E., and T. H. Lee. 2015. "Why Strategy Matters Now." *New England Journal of Medicine* 372 (18): 1681–4.
- Roland, M., and R. A. Dudley. 2015. "How Financial and Reputational Incentives Can Be Used to Improve Medical Care." *Health Services Research* 50 (6 Pt 2): 2090–115.
- Ryan, A. M., and M. J. Press. 2014. "Value-Based Payment for Physicians in Medicare: Small Step or Giant Leap?" *Annals of Internal Medicine* 160 (8): 565–6.
- Schlesinger, M., R. Grob, and D. Shaller. 2015. "Using Patient-Reported Information to Improve Clinical Practice." *Health Services Research* 50 (6 Pt 2): 2116–54.
- Steinbrook, R. 2015. "The Repeal of Medicare's Sustainable Growth Rate for Physician Payment." *Journal of the American Medical Association* 313 (20): 2025–6.
- VanLare, J. M., and P. H. Conway. 2012. "Value-Based Purchasing—National Programs to Move from Volume to Value." *New England Journal of Medicine* 367 (4): 292–5.
- Werner, R. M., and R. A. Dudley. 2012. "Medicare's New Hospital Value-Based Purchasing Program Is Likely to Have Only a Small Impact on Hospital Payments." *Health Affairs (Millwood)* 31 (9): 1932–40.
- Wynia, M. K. 2009. "The Risks of Rewards in Health Care: How Pay-for-Performance Could Threaten, or Bolster, Medical Professionalism." *Journal of General Internal Medicine* 24 (7): 884–7.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.