

Impact of Incentives to Improve Care for Primary Care Patients

LISE RYBOWSKI, MBA; DALE SHALLER, MPA; SUSAN EDGMAN-LEVITAN, PA;
AND PAUL D. CLEARY, PHD

ABSTRACT

OBJECTIVES: Whether incentive programs lead to more effective improvement activities by medical practices remains an open question. We assessed the impact of a pay-for-performance program designed to improve patients' experiences with primary care and identified the factors that influenced the outcomes.

STUDY DESIGN AND METHODS: We conducted telephone interviews with clinical and administrative leaders of 8 physician practices that had been identified by a major health insurance plan as needing to improve on their past scores on a patient experience survey.

RESULTS: The financial incentive had almost no effect on the priorities or activities of the medical practices. Few practice leaders were aware of the incentive. All were familiar with local efforts to improve patients' experiences and some maintained that focus over time, but not because of the ongoing system-level initiative. The effectiveness of the incentive program was limited by the program design, the improvement goals, the timing of progress measurement, and a lack of ongoing communication about the program's purpose, available support, and goals. Medical practice leaders' understanding of the program and its goals was limited; providers infrequently used the free educational programs available; and other initiatives may have taken priority.

CONCLUSIONS: These findings suggest that if system-level incentives are to improve care quality, they must be designed carefully to reach the audience responsible for improving care, to motivate organizational change, and to support ongoing communications with practice leaders.

Both public and private payers have been expanding the use of pay-for-performance (P4P) programs that encourage improvements in primary care by linking financial rewards to evidence of higher quality. These programs have evolved in response to concerns that many weaknesses in the healthcare system are a result of the way providers are compensated.

Whether P4P programs lead to improvements in the quality of healthcare remains an open question. There are relatively little data on the influence of such programs^{1,2} and the available evidence on their effectiveness is mixed.³ Studies that show an impact have found modest effects⁴; these studies have focused on measures of clinical process and intermediate outcomes.⁵ Given the increasing attention to patient experience measures in value-based purchasing programs, such as the CMS Medicare Shared Savings Program,⁶ it is important to understand the effectiveness of P4P strategies in driving improvements in patient experience.

The purpose of this study was to assess whether and how an incentive program for a large provider network to improve primary care patients' experiences affected the priorities and activities of medical practices. In this article, we describe the program and use the framework proposed by Van Herck and colleagues to assess several factors that might have affected its results.⁷ The review of 128 evaluation studies by Van Herck and colleagues concluded that the impact of P4P programs can be influenced by the program's context as well as various design choices, including the ways in which quality is measured, the quality goals and target, the nature of the incentives, how the program is implemented and communicated, and how the effects are evaluated. This framework offers useful insights into the factors that may have undermined or facilitated the effectiveness of the P4P program in driving improvements in patient experience.

Setting

In 2004, a large healthcare network negotiating with a major health insurance plan in its market⁸ agreed to a 5-year P4P program that would financially reward practices for attaining specified goals. An improvement in patients' experiences with primary care was one of the performance goals for the second phase of the contract (2007-2009). The incentive was that part of a "revenue withhold" would be

returned to the network if lower-performing primary care practices improved their scores on a standardized patient experience survey adapted from the Consumer Assessment of Healthcare Providers and Systems (CAHPS) Clinician & Group Survey.⁹ The program included 96 primary care practice sites in the network that scored below the statewide mean on 1 or more measures of patient experience. The total value of the withhold tied to patient experience was about \$5 million.

To help the low-performing practices improve patient experience, the network contracted with an internal center with expertise in primary care improvement. The center supported the practices in developing priorities and improvement plans and offered free educational and consulting services designed to help the practices improve the domains of patient experience addressed by the survey.

Under the terms of the P4P contract, the network was required to submit improvement plans for 95% of low-performing practices. In the third year, the practices were expected to achieve specified improvements in performance. The percentage of the withhold that would be available to the practices ranged from 25% to 100%, depending on the percentage of practices that met their target based on the 2009 survey results. If fewer than 55% of practices met the target, none of the withheld revenue would be returned.

Methods

In 2008, 2 of the authors (LR and DS) conducted in-person and telephone interviews with several individuals involved in the P4P program from the health plan, the provider network, the internal consulting center, and an independent organization that gathers and reports the survey results. These interviews elicited information about the P4P program as well as perspectives on the program's development, purpose, and feasibility.

The 96 physician practices that met the criteria for participation in the P4P program were affiliated with 16 large organizations that manage the practices. To select potential interviewees for this study, we first identified the 2 survey topics chosen as improvement priorities by the most practices: doctor-patient communication and office staff courtesy and respect. Our sample consisted of the 31 practices that chose either of the 2 topics; these practices represented 11 of the large physician organizations.

The 31 practices were divided into 2 groups: those that met the target (16) and those that did not (15). The original goal was to interview 2 practices that met the goal and 2 that did not for each of the 2 topics. However, it was not possible to identify practices in all 4 categories because all of the practices that focused on communication met their target. Thus, we selected practices that represented a mix of topics (ie, both communication and office staff), a mix of results for the office staff measure, and a mix of large physician organizations.

The study protocol was approved by the Yale Human Investigations Committee. All subjects knew that we were conducting research and agreed to be interviewed.

In 2011, we invited 17 practices by e-mail and phone to partici-

pate in 30- to 60-minute telephone interviews. If a practice declined to participate, another practice was selected and contacted. In total, the authors interviewed clinical and administrative leaders of 8 medical practices, 1 of which included multiple sites. These sites represented approximately one-third of the medical practices focusing on the 2 topics: 11 of the 31 practice sites and 5 of the 16 large physician organizations.

Results

The interviews with the clinical and administrative practice leaders revealed a low level of awareness of the P4P program. Only 1 of the interviewees knew that the practice had a financial incentive to improve patient experience scores; none was aware that the network did not receive the withheld dollars tied to patient experience because slightly fewer than 55% of the participating practices had met their performance target.

Moreover, the practice leaders were not aware that the goal of improving patient experience was part of an ongoing network initiative. While they knew about the network's initial efforts to focus their attention on patient experience, those that maintained that focus over time did not characterize their efforts as part of a system-wide initiative to achieve specified goals.

While all the practices were aware of the public report of patient experience survey results, several had never looked at the report and almost none had used it. To assess their performance, they relied on their own survey results or those collected for their practice by the provider organizations with which they were affiliated. The differences between the publicly reported measures and scores and those from other sources caused confusion and frustration.

The practice leaders also varied significantly in their awareness, understanding, and use of the free educational and consulting services. While some took full advantage of the offerings, others were either unaware of the services or uninterested. Also, even though several practice leaders indicated that doctors or staff had participated in and benefited from at least 1 educational event, not all of them were clear on what organization sponsored the events or that the services were related to the incentive to improve patient experience.

Because of the requirement to produce improvement plans, the P4P program initially succeeded in capturing the practices' attention. However, the long-term incentive generally did not lead to an ongoing focus on patient experience. While a few practices continued to plan and implement improvement strategies, most turned their attention to other challenges, noting that patient experience was only one of many issues that require their attention and resources.

Factors Affecting Program Success

Categorizing the interview comments according to the framework of Van Herck and colleagues⁷ suggests why the P4P program failed to bring about a sustained effort to improve patient experience.

The External Context for a P4P Program

The network-level incentive to improve patient experience took

place in a context that placed conflicting demands on practices. First, this program coincided with a statewide health reform effort that resulted in insurance coverage for most residents. For many practices, the subsequent pressure to meet the demand for primary care services overwhelmed their ability to focus on other needs. All of the practices noted that improving patients' access to care and information had been one of their highest priorities in recent years. While access to care is an element of patient-centered care, it was not the domain that these practices had committed to improve under the P4P program. Thus, to some extent, the practices' commitment to improving one aspect of patient experience was competing with the urgent need to improve another aspect.

Constant competition for the attention of administrative and clinical leaders was also a key part of the external context. Within the network, the goal to improve patient experience was just 1 of several P4P goals. When measures and goals change from year to year, it is a significant challenge for healthcare leaders to maintain a focus on any improvement target over time.^{5,10} In addition, other health plans and payers were seeking improvements in specific areas. Health plans in the state had been implementing performance incentives for several years: a 2004 survey found that 89% of the medical groups had a P4P incentive in at least 1 commercial plan contract. Over a third of the surveyed groups reported that their incentives were tied to their performance on patient satisfaction surveys.¹¹

Finally, organizational identity and affiliation contributed to the context for the participating medical practices.¹² None of the interviewees said that their improvement work was focused on a goal shared with primary care practices across the provider network. To the extent that they referred to anything other than their own practices, it was their affiliation with a large health system or medical group. However, the goals of the network's P4P contract were not always aligned with the goals and programs of these other, more proximate organizations.

Friedberg et al found that a majority of medical groups in the study state were working to improve patient experience: 61% reported groupwide improvement efforts and 22% were focused on improving care from low-scoring physicians or practice sites.¹³ Those groups were more likely than other groups to have some financial incentive to improve. Thus, it is likely that at least some of the practices affected by the network-level program were also participating in a group-level program. However, the improvement targets and strategies for the medical groups were not identical. Although both programs focused on communication, for example, the improvement initiatives at the group level did not emphasize doctor-patient communication; they were primarily focused on organizational factors, such as redesigning office work flow, training nonclinicians, using electronic health records, and reassigning staff responsibilities.¹³

The Quality Measure

The measures used to assess practice performance are a key element of a P4P program. While consumers and healthcare providers

have had access to reports with results of the CAHPS Clinician & Group Survey since 2006, the interviews revealed that not everyone accepted the survey scores as an accurate reflection of quality. Some practice leaders noted that a small subset of their physicians were skeptical that quality as measured by the patient experience survey was even a problem. Others accepted the measures but expressed defensiveness about their performance, claiming that the results were not adequately adjusted to reflect differences in patient populations. This lack of trust in the measures may have undermined efforts to improve quality.¹²

Quality Goals and Targets

One positive element of this program was that its focus on practices with a baseline performance below the 2007 mean meant that all participants had room to improve. Because of the timing of the surveys, however, it was not possible to assess the degree of improvement until the results of the 2009 survey had been analyzed. Practices that initiated changes to improve patient experiences in 2008 had to wait until the spring of 2010 to find out whether they achieved their targets on the 2009 survey.

More frequent assessments and payments (eg, quarterly bonuses) might have served as an important reminder of the P4P program and its goals.¹⁰ Some practices tried to set interim goals by examining regular feedback based on other patient experience surveys, but found that they could not always relate the measures in those surveys to the measures in the statewide survey.

Nature of the Incentive

Van Herck et al found that the incentive size for many P4P programs in the United States is low (estimated at 1% to 2% of income) and that there was no relationship between incentive size and effect.⁷ The issue for this program was not the size of the incentive but the lack of clarity about how much money was at stake for a practice or an individual physician. The importance of targeting an incentive at the individual and/or team level is well established.⁷ However, the provider network allowed the large organizations that manage groups of practices to determine the flow of funds from P4P activities. Some kept all P4P money they earned to support the organization while others shared the funds with the practices. Because of variation in the way incentives were distributed, no practice knew what the incentive was worth to them, and only some practices offered incentives to individual physicians.

Furthermore, although the unit of analysis for determining progress was the primary care practice, whether an incentive was earned depended on the combined performance of a group of low-performing practices, which was a subset of all primary care practices in the network. Practices that succeeded in reaching their own target did not get anything if the entire group did not make its overall target. It is possible that this "shared benefit" design led some practices to assume that they did not have to put much effort into improving

on the selected measure because everyone else would do what was needed to gain the incentive.

Implementing and Communicating the Program

Another factor that may have undermined the effect of the P4P program was that once medical practices had submitted improvement plans, they were not required to do anything else, including developing the capacity to manage and conduct improvement work.¹⁰ Among the interviewed practices, none had a dedicated quality improvement team and only a few had someone in a leadership role with expertise in quality improvement methods.

The limited awareness of this program points to a failure to communicate with the participating providers, which has been associated with the lack of an effect in other P4P programs.⁷ The network leadership did not continue communicating and emphasizing the importance of the program's goals. With the exception of 2 newsletter articles, the practices did not receive regular communications about this program to reinforce its goals and rationale, to explain how the program worked, or to hold practice leadership accountable. The center under contract to support the practices had contact with the practices, but the staff did not have access to established communication systems, nor were they allowed to contact the practice administrators directly. The provider network required that they communicate with the practices through the network's account managers.

Individuals from about half of the study practices indicated that someone in the practice took advantage of the training and other resources provided by the center tasked with supporting the practices. However, there were some misunderstandings and confusion about the nature and scope of the resources. Several interviewees seemed unsure about the types of resources that were available and questioned whether the resources were applicable to the issues they needed to address. They also expressed confusion about who was offering the resources—even if they had already attended a training session—and whether they could or should take advantage of the offerings, which appeared to some to be offered by a rival hospital system or to be redundant with other services.

How the Effects Are Evaluated

As noted above, there was a significant gap in time between the practices' initial efforts and the reporting of results from the 2009 survey, which made it difficult for the practices to gauge whether they were making progress. Furthermore, the results from the 2007 survey were not available until April 2008, and the improvement plans had to be finalized by September 2008. The time between the implementation of those plans and the follow-up measurement in the fall of 2009 may not have been sufficient to assess the impact on patient experience.

Discussion

This study points to several factors that have the potential to limit the effectiveness of incentives. The first is the flow of money (ie, whether the financial reward actually accrues to the organizational units or individuals responsible for improvements in care). The second factor is the nature of communication among the different levels of a large healthcare organization. In the absence of clear and ongoing communication about the incentive, there was no shared understanding of the goals or who stood to gain or lose, thus limiting the effect of the incentive. Insufficient communication also meant that practices were not aware of and made little use of the educational and consulting support available to them.

This study also suggests that P4P programs may benefit from some coordination with other similar quality improvement programs in the region—or at least efforts to ensure that the goals will not conflict with those of other initiatives. In this case, the primacy of improving access to primary care may have taken precedence over other patient experience goals.

Limitations

Two potential limitations of this study were the small number of practices that we studied and the timing of the interviews. The responses of the selected practices, about 10% of the total number of practices in the P4P program, may not be representative of all other practices. Timing may be an issue because the interviews with practice leaders took place more than a year after the pay-for-performance program was initiated. In many cases, the leaders had moved on to new goals and did not necessarily remember what they had done to meet their improvement target, when they had done it, or why. Nevertheless, the results reported here support our conclusion that most practices were not taking a long-term approach to addressing the performance issues associated with the incentive.

Conclusions

The use of system-level incentives as described in this study was not an effective strategy for encouraging improvements in patient experience with care. Future efforts to motivate improvement across a health system may benefit from a more direct connection of the financial reward or penalty to those responsible for improving care, clear and ongoing communications with clinical and administrative leaders, and better coordination among the organizations that set improvement goals for physician practices.

Author Affiliations: The Severyn Group (LR), Ashburn, VA; Shaller Consulting Group (DS), Stillwater, MN; John D. Stoeckle Center for Primary Care Innovation, Massachusetts General Hospital (SEL), Boston; Yale School of Public Health (PDC), New Haven, CT.

Source of Funding: Work on this project was supported by a cooperative agreement from the Agency for Healthcare Research and Quality (#U18HS016978).

Author Disclosures: The authors report no relationship or financial

interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (LR, DS, PDC, SE-L); acquisition of data (LR, DS, SE-L); analysis and interpretation of data (LR, DS); drafting of the manuscript (LR, SE-L); critical revision of the manuscript for important intellectual content (DS, PDC); provision of patients or study materials (SE-L); obtaining funding (PDC); administrative, technical, or logistic support (SE-L); and supervision (SE-L).

Address correspondence to: Lise Rybowski, MBA, The Severyn Group, 21121 Stonecrop Pl, Ashburn, VA 20147. E-mail: lise@severyn-group.com.

REFERENCES

1. Mullen KJ, Frank RG, Rosenthal MB. Can you get what you pay for? pay-for-performance and the quality of healthcare providers. *RAND J Econ*. 2010;41(1):64-91.
2. Rosenthal MB, Frank RG. What is the empirical basis for paying for quality in health care? *Med Care Res Rev*. 2006;63(2):135-157. Review.
3. Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev*. 2011;(7):CD009255.
4. Scott A, Sivey P, Ait Quakrim D, et al. The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Syst Rev*. 2011;9:DC008451.
5. Rosenthal MB, Landon BE, Howitt K, Song HR, Epstein AM. Climbing up the pay-for-performance learning curve: where are the early adopters now? *Health Aff (Millwood)*. 2007;26(6):1674-1682.
6. CMS Shared Savings Program: quality measures and performance standards. CMS website. http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/shared-savingsprogram/Quality_Measures_Standards.html. Updated March 2, 2015. Accessed May 12, 2015.
7. Van Herck P, De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W. Systematic review: effects, design choices, and context of pay-for-performance in health care. *BMC Health Serv Res*. 2010;10:247.
8. Competition in Health Insurance: A Comprehensive Study of U.S. Markets 2007 Update. Chicago, IL: American Medical Association; 2007.
9. Dyer N, Sorra JS, Smith SA, Cleary PD, Hays RD. Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician and Group Adult Visit Survey. *Med Care*. 2012;50 suppl:S28-S34.
10. Davies E, Cleary PD. Hearing the patient's voice? factors affecting the use of patient survey data in quality improvement. *Qual Saf Health Care*. 2005;14(6):428-432.
11. Mehrotra A, Pearson SD, Coltin KL, et al. The response of physician groups to P4P incentives. *Am J Manag Care*. 2007;13(5):249-255.
12. Nembhard IM, Alexander JA, Hoff TJ, Ramanujam R. Why does the quality of health care continue to lag? insights from management research. *Acad Manage Perspect*. 2009;23(1):24-42.
13. Friedberg MW, Steelfisher GK, Karp M, Schneider EC. Physician groups' use of data from patient experience surveys. *J Gen Intern Med*. 2011;26(5):498-504.