Organizational Processes and Quality

Putting Smart Money to Work for Quality Improvement

Nancy Dean Beaulieu and Dennis R. Horrigan

Objective. To investigate the effects of paying physicians for performance on quality measures of diabetes care when combined with other care management tools.

Data Sources/Study Setting. In 2001, a managed care organization in upstate New York designed and implemented a pilot program to financially reward doctors for the quality of care delivered to diabetic patients. In addition to paying a performance bonus, physicians were also supplied with a diabetic registry and met in groups to discuss progress in meeting goals for diabetic care. Primary data on diabetes care at the patient level were collected from each physician during the 8-month period, April 2001–January 2002.

Study Design. Physicians were scored on individual process and outcome measures of diabetes care on three separate occasions; these individual scores were combined into a composite score on which the financial reward was allocated. The study design is pre/post for the patients whose physicians participated in the performance pay program. The control group is a large sample of the health plan's diabetic members.

Data Collection. Data on patient outcomes were self-reported by physicians participating in the study. These data were audited with spot checks of medical charts. Data for the control group were collected as part of the health plan's annual HEDIS data collection.

Principal Findings. Physicians and patients achieved significant improvement on five out of six process measures, and on two out of three outcome measures (HbA1c control and LDL control). Thirteen out of 21 physicians improved their average composite score enough to earn some level of financial reward. Of the eight physicians not receiving any of the three levels of reward, six improved their composite scores.

Conclusions. Financial incentives for physicians, bundled with other care management tools, led to improvement on objectively measured quality of care for diabetic patients. Self-selection by physicians into the pay pilot and the small sample size of participating physicians limit the generalizability of the results.

Key Words. Physician incentives, quality improvement, pay for performance

The influential 2001 report by the Institute of Medicine (IOM) made public an analysis of quality problems in the U.S. health care system and put forth

recommendations for quality improvement (Institute of Medicine 2001). One of the report's key recommendations is a fundamental change in payment methodologies to reward quality. In a recent article by Rosenthal et al. (2004), surveying the use of a variety of incentive systems that reward quality, the authors conclude by noting that no systematic evaluations of the intended or unintended consequences of such incentive systems have been conducted. This article begins to fill that gap by reporting on a demonstration project in which Independent Health (IH), a managed care plan in upstate New York, paid some of its doctors for the quality of care received by their diabetic patients. The operational details and the lessons learned from this demonstration project offer key insights into how one might implement an important element of the IOMs quality agenda, and the early steps we might take toward a system fundamentally changed to reward and deliver quality health care.

STATE OF THE ART IN FINANCIAL INCENTIVES FOR QUALITY

The health care system is connected by an extensive array of contractual agreements among employers, subscribers, health plans, physicians, and hospitals. These agreements generally contain provisions for coverage, access, and payment, but do not have specific performance requirements for clinical quality. Over the past decade, health plans have used a variety of approaches to provide financial incentives for quality improvement. These include quality bonuses, return of compensation at risk (withholds), performance fee schedules, and reimbursement for care planning. However, most quality incentive programs in use today are evidence of tinkering around the edges of traditional payment systems.

These traditional incentive programs have a few characteristics in common. First, they are typically grafted on to existing payment policies and, with the possible exception of the return of withholds, often do not constitute significant financial incentives for physicians. Second, in practice, many of these "quality incentives" are actually tied to performance on utilization measures (e.g., emergency room visits). Third, contracting physicians are typically not involved in the design of these incentive programs. Fourth, these

Address correspondence to Nancy Dean Beaulieu, Ph.D., Assistant Professor, Harvard Business School, Mellon Hall A4-5, Soldiers Field, Boston, MA 02163. Dennis R. Horrigan, M.S., is Vice President with Managed Care Development, Independent Health Association, NY.

incentives are not often bundled with other activities supporting quality improvement.

Traditional incentive programs are unlikely to result in substantial and permanent quality improvement for at least two reasons. First, physicians often view these incentive programs with skepticism. Their “add-on” nature gives the impression that quality is much less important than cost-savings or productivity. Furthermore, physicians sometimes view the chosen quality measures as relatively uninformative indicators of overall quality of care (e.g., rates of mammography) and illegitimate because of measurement issues (e.g., problems with risk adjustment). Compounding these problems is the fact that these incentive programs typically operate in the shadow of an adversarial health plan–physician bargaining relationship. This leads physicians to be suspicious of health plans’ true objectives and intentions with regards to quality incentive programs.

Second, traditional quality incentive programs are infrequently paired with complementary quality improvement activities (like training, infrastructure investment, etc.). The chronic care model developed by Edward Wagner and colleagues highlights the critical importance of decision-support, information systems, and delivery system design for improving care to patients with chronic health care problems (Bodenheimer, Wagner, and Grumbach 2000b, 2002a). In addition to these systems, expertise is needed to analyze data, hypothesize causes, and devise improvement strategies. Most physician offices lack these systems and expertise. Thus, the mere creation of financial incentives will not close the quality chasm; organizations must concurrently develop supporting infrastructure and develop new capabilities for continuous improvement.

WESTERN NEW YORK PHYSICIAN INCENTIVE PROJECT

In 2002, IH, a managed care organization in upstate New York, initiated a quality incentive project for primary care physicians. The goals of this project were threefold: (1) to improve chronic care treatment for diabetic members, (2) to explore the effectiveness of financial incentives in improving care for this patient population, and (3) to promote the development of office-based systems of care.

This quality incentive project built on the health plan’s experience with a more traditional incentive program put in place in 1990. This traditional program provided financial rewards to primary care physicians whose patients

Table 1: Activities Prescribed by the Incentive Plan

<i>Health Plan Activities (1)</i>	<i>Physician Activities (2)</i>
1. Provide physicians with a clinical practice recommendation for diabetic care	1. Integrate clinical practice recommendations into workflow
2. Provide physicians with a chart assessment tool	2. Conduct chart review and collect data
3. Provide physicians with a registry or list of patients with the diagnosis of diabetes	3. Share results with other practices
4. Analyze and report data to each physician three times during the year	4. Organize office staff to engage diabetic patients in best practice
5. Organize collaborative meetings to assist the physician in learning from others	5. Provide follow-up and case findings for patients
6. Provide patient education materials	
7. Educate and train physician office staff	
8. Regular communication between physicians, staff, and the health plan's physician account executive	

met specified targets in areas of access, patient satisfaction, and preventive health care. The diabetic improvement initiative was designed to establish a quality measure for a chronic condition that could be eventually integrated into the traditional program.

From the start, the project team at IH believed that in order to achieve lasting quality improvement in patient care and outcomes, the incentives would need to influence care delivery processes on-site at the physician's office. As a consequence, the team bundled financial incentives with other interventions intended to (1) facilitate the establishment of new routines in the physician's office, (2) educate physicians, and (3) increase communication between physicians and the health plan, among physicians, and between physicians and their patients. Column 1 of Table 1 lists specific complementary activities undertaken by IH to instigate real behavioral change.

Design of the Reward Program

The design and implementation of incentive programs to financially reward quality of care present many conceptual and operational challenges for physicians and health plans. Conceptually, physicians view quality as emanating from training and licensure, and identify inadequate reimbursement as a major contributor to suboptimal quality. In contrast, health plans seek to uncover underuse, overuse, and misuse of services as the sources of poor quality

(themes from the IOM reports) and to utilize marketplace solutions to stimulate quality improvement.

Health plans and physicians also have divergent opinions about how to fund incentive programs. Health plans expect revenues for incentive payments to be generated by reducing unnecessary variation or by reducing payment to poor performing physicians. Physicians view this approach as disingenuous and suspect that the funds will never be available for disbursement. Physicians also are reluctant to support incentive programs that are designed to be budget-neutral and are funded by limiting payment to underperforming physicians.

Even when agreement is reached on these conceptual issues, health plans and providers must achieve consensus on the more operational issues such as what conditions need improvement, corresponding measures of quality, and target improvement levels to be achieved. For example, rewarding physicians for diabetic A1c control at 7.5 or less is a simple and clear goal and, while a stretch for many offices, is a well-accepted measure of quality.

With these thoughts in mind, an IH team created an incentive program that was purposefully disconnected from the ordinary business of reimbursement for health care services delivered to its members. Diabetes was selected because of the documented gaps in treatment, the expected impact of guideline compliance on medical outcomes, and the availability of credible quality measures. The financial incentive took the form of a quality bonus to be paid on an annual basis to physicians scoring above a predetermined target on a composite performance index.

The IH team selected both process and outcome measures of quality to measure performance and combined these into a composite score that was computed at both the individual patient level and the practice level. The specific measures and their weighting in the composite score are listed in Table 2. The composite elements were chosen to conform to evidence-based recommendations for diabetic care derived from the American Diabetes Association clinical guideline.

Outcome or physiologic measures received greater weight in the composite than process measures; this design signalled that the ultimate goal of the program was to improve patient health. Note that all of the process measures chosen were designed to monitor some aspect of the progression of the patient's disease. Interventions such as nutrition counseling, adoption of an exercise program, and compliance with a drug regimen would be likely to improve outcomes. However, patient engagement in these activities is dependent more upon the patient's preferences and less upon the potential

Table 2: Composite Performance Measures

<i>Measures</i>	<i>Weighting Used</i>	<i>Measures</i>	<i>Weighting Used</i>
1. One A1c test during the 12-month measurement period	0.5	6. One dilated retinal exam during the measurement period	1.0
2. A second A1c test completed 4–6 months after test 1 and within the 12-month period	0.5	7. Lower extremity foot exam documented in chart (monofilament/tuning fork)	1.0
3. A1c value less than or equal to 7.5%	2.0	8. Annual urine microalbumin/creatinine test	0.5
4. One LDL test during the measurement period	0.5	9. Blood pressure $\leq$ 130/80 during measurement period	2.0
5. LDL test during measurement period $\leq$ 100 mg/dl	2.0	Totals for care	10

actions of the physician, and performance on these measures would be difficult to monitor and verify.

There were two different ways by which a physician could earn a financial reward: by meeting one of two benchmarks, or by posting a 50 percent improvement in the composite score. The composite score of 6.86 reflects a level of best practice for diabetic care and the composite score of 6.23 reflects a level that, while not achieving best practice, requires achievement of one level of physiologic control. The 50 percent improvement target was designed to provide a positive incentive for those practices that achieved substantial improvement but did not meet one of the two performance targets. The performance targets chosen by IH were unlikely to be viewed by physicians as arbitrary since they were computed for a hypothetical physician who employed “best practice” according to the clinical guidelines established by the American Diabetes Association. For reasons of professionalism, physicians whose performance fell below the targets may have been strongly motivated to improve.

There are obviously different ways to achieve the same composite score. From an economic efficiency point of view, this is desirable because it permits physicians to make changes that are least costly to their individual practices. However, measuring physician performance as the average over all patients also permits the physician to focus efforts on some patients (e.g., the easier or healthier patients) and not others (e.g., the more difficult or sicker patients). Note also that the combination of process and outcome measures attenuates the need for precise risk adjustment of the average composite score across

Table 3: Schedule of Financial Rewards

<i>Score</i>	<i>Percentage of Incentive</i>	<i>Medicare</i>	<i>Commercial</i>
Composite score ≥ 6.86	100	\$3.00 PMPM	\$.75 PMPM
Composite score ≥ 6.23	50	\$1.50 PMPM	\$.37 PMPM
50% improvement in composite score and composite score ≤ 6.23	25	\$.75 PMPM	\$.18 PMPM

PMPM, per member per month.

practices. For those physicians with patients who have advanced stages of diabetes (and for whom it might be more difficult to achieve performance gains on outcome measures), it is still possible to score highly on the process measures.

Physicians who met the targets or demonstrated significant improvement received a per member per month (PMPM) bonus paid on the total number of the health plan's members in the physician's panel. The levels of financial reward associated with the performance targets are listed in Table 3; the different reward levels for commercial and Medicare patients reflect differences in overall reimbursement for these two product lines. Physicians achieving a score of 6.86 or greater would receive an incentive payment equivalent to a 12 percent increase in PMPM reimbursement. This is true for both fee-for-service and capitated physicians. Actual payments ranged from \$3,000 to \$12,000 based on panel size and performance.

Implementation

In the fall of 2001, 34 primary care physicians from the northern part of Western New York were contacted and informed about the health plan's desire to improve the care of patients with diabetes and that additional compensation would be made to physicians who reached target levels of improvement. Twenty-two of the 34 physicians enrolled in the program and signed an agreement that specified the responsibility of the physician to engage in the activities listed in column 2 of Table 1. During the study, one physician retired and the physician who bought the practice elected not to continue in the incentive program, leaving us with a final sample of 21 physicians and 624 diabetic patients. Twelve physicians elected not to participate; their rationale ranged from not wanting to invest the time to simply ignoring the opportunity.

Currently, most clinical offices do not have an efficient method to identify their patients by diagnosis. Based primarily on claims analysis, IH created a registry which included the names of all diabetic patients assigned to each primary care physician in all lines of business including Commercial, Medicare, and Medicaid. The health plan produced a paper registry for each physician; the average number of IH members in each physicians practice was 800 and the average number of diabetic patients per physician was 32. As performance data were accumulated for each patient, it was added to the registry. This registry contained data for all diabetic patients in the practice and promoted practice-based population health management.

In order to dispense rewards, the IH team faced the challenge of collecting detailed performance data from each of the participating physicians three times during the year. They designed an instrument for the physician to record the performance measures for each diabetic patient. This instrument enabled the physician to score performance for each patient and for all diabetics in the practice; it also facilitated comparisons to other practices.

The data collection approach required physicians to do their own chart assessments for individual patients and to self-report these scores; the health plan contributed data analysis, auditing, and reporting. This division of duties was purposefully chosen. The IH program design team thought that having the physicians and their staff immediately involved in the recording and collection of data could generate multiple benefits. First, this method would provide physicians and their staff with a clear understanding of the program and a constant reminder of the performance measures and absolute targets. Second, a logical place to keep track of the performance data at the individual level is in each patient's medical record. Thus, a checklist in the patient's chart would serve as a timely reminder to the physician to schedule the appropriate tests and to engage the patient in conversation about the status of their disease and the patient's progress. Third, the presence of longitudinal performance data in the medical record provided the opportunity and the prompt for physicians to examine trends over time, to hypothesize about determinants of these trends, and to develop strategies for influencing these trends.

Performance data on individual patients were collected three times during the year (baseline, interim, and final) and the reward, when merited, was granted at the end of the year. The baseline score was computed in March of 2002 based on care delivered during the previous 12 months; the interim score was computed in August of 2002 again based on care delivered during the previous 12 months. The final score was computed in January of 2003 based on care delivered during 2002. This schedule enabled physicians to assess

patient status three times and to implement two cycles of improvement in the office.

At the end of each measurement period, health plan personnel would gather together a few doctors at a time to review the results. Each physician was provided with his or her own data, along with blinded data for other participating physicians. These meetings facilitated discussions among the physicians about how to improve care for specific patients and, in general, the types of processes used in the office that facilitated improvement across all patients.

RESULTS

Results—Patients

As shown in Table 4, the diabetic care initiative demonstrated significant improvement on five out of six process measures over the 8-month study period. The sample for these measures is 476 members for whom we had complete performance data and who were enrolled continuously from the baseline measurement period.

For reference as a comparison group, we have included in Table 4 IHs HEDIS performance scores on five of the process measures (the HEDIS set did not include the foot exam measure). The HEDIS performance scores are based on a sample of 600 of IHs diabetic members for each product. Changes

Table 4: Performance on Process Measures

<i>Measure</i>	<i>Population</i>	<i>Baseline Performance 2001</i>	<i>Final Performance 2002</i>	<i>Change in Performance (2002–2001)</i>
HbA1c test (1)	Pilot study patients	80.3	83.4	3.1
HbA1c test (2)	Pilot study patients	52.7	78.2	25.5*
HbA1c test (1)	IH diabetic members	87.1	86.5	– 0.6
Lipids test	Pilot study patients	68.5	86.8	18.3*
Lipids test	IH diabetic members	88.8	91.0	2.2
Diabetic retinal exam	Pilot study patients	37.4	63.0	25.6*
Diabetic retinal exam	IH diabetic members	50.3	51.5	1.2
Nephropathy test	Pilot study patients	41.8	78.8	37.0*
Nephropathy test	IH diabetic members	48.5	51.6	3.1
Foot exam	Pilot study patients	32.6	77.9	45.4*

*Baseline performance significantly different from final performance, $p < .001$.

in these HEDIS performance scores during the period of the demonstration project control for other activities occurring within the health plan and within the larger health care community that might have independently led to improved performance on the process measures. Improved performance in the study group is an order of magnitude greater than the improved performance in the control group (the HEDIS scores). In addition, the performance on the process measures at baseline is in all cases below the performance of the control group at baseline; this comparison suggests that the incentive program may have been successful in generating improvement among patients with the lowest guideline compliance.

It is noteworthy that, except for the diabetic retinal exam, the magnitude of the increase in the performance score (measured both in absolute and percentage terms) on these process measures is inversely related to the baseline score. This pattern suggests that the cost or difficulty of improvement is increasing in the performance score and that continued improvement might require alternative approaches and/or larger financial incentives.

As shown in Table 5, the incentive program had similarly positive effects on outcome measures. The number and percentage of patients whose A1c, cholesterol, and blood pressure were under control rose during the study period. IH implemented stricter thresholds to define control than those used in the HEDIS quality measures. Rates of control in the study population were also computed using the HEDIS definitions to facilitate a comparison between changes in performance over time in the study population and IHs diabetic population (the control group).¹ During the time period of the study, the HEDIS scores for the control group show that the rate of A1c control changed very little and the rate of lipid control actually declined. In contrast, control

Table 5: Performance on Outcome Measures

	<i>Population</i>	<i>Base 2001</i>	<i>Final 2002</i>	<i>Change Base to Final</i>
HbA1c $\leq$ 7.5	Pilot study patients	53.8	56.1	2.3
HbA1c $\leq$ 9.5	Pilot study patients	61.8	75.6	13.9*
HbA1c $\leq$ 9.5	IH diabetic members	72.4	74.2	1.8
LDL $\leq$ 100	Pilot study patients	38.4	48.9	10.5*
LDL $\leq$ 130	Pilot study patients	46.0	69.5	23.5*
LDL $\leq$ 130	IH diabetic members	60.4	59.7	- 0.8
BP $\leq$ 130/80	Pilot study patients	60.3	66.6	6.3**

*Baseline performance significantly different from final performance, $p < .001$.

**Baseline performance significantly different from final performance, $p < .05$.

rates for A1c and LDL (under the HEDIS definition of control) increased substantially in the study population. Blood pressure control is not included in the HEDIS measurement set for diabetes; consequently, we have no control group against which to judge the effectiveness of the program intervention on this outcome measure. Absolute gains in blood pressure control among the study population were modest but still significant.

Results—Doctors

The average of the physicians' composite scores increased 48 percent from 4.05 at baseline to 5.99 at the end of the project. Figure 1 shows the distribution of the average composite scores for physicians at the end of each measurement period. To put these changes in perspective, we compare the baseline score for physicians in this study to the baseline score of a large sample of the health plan's primary care physicians. In the late spring of 2003, approximately 300 primary care physicians were asked to conduct a self-assessment of their care for a sample of diabetic patients in their practice ($n = 15$ for each physician). Diabetic care was assessed using the same format used in pilot the study. Participating physicians received a \$0.30 PMPM payment for completing the survey; they received no payment for their performance on the measures. Eighty-five percent of these physicians completed the assessment. IH subsequently scored these physicians using the same 10-point composite performance index used in the study; the average score in this sample of physicians was 4.4.

Figure 1: Distribution of Average Physician Composite Scores over Time

Table 6: Changes in Average Composite Score by Level of Reward

<i>Level of Reward (%)</i>	<i>Number of Physicians</i>	<i>Average Baseline Composite Score</i>	<i>Average Final Composite Score</i>
100	8	5.55	7.64
50	2	2.82	6.61
25	3	1.76	4.70
0	8	4.29	4.58

As shown in Table 6, 13 of 21 physicians earned a financial reward. Eight physicians finished the year with a composite score equal to or greater than 6.86; two physicians cleared the lower hurdle of 6.23; and three physicians achieved at least a 50 percent improvement in their composite scores though their final scores were less than 6.23. Of the eight physicians not receiving any of the three levels of reward, six improved their composite scores.

One of the major theoretical concerns in paying for quality relates to the problem economists refer to as distortion. A pay for performance scheme may induce effort on the rewarded measures of performance and may reduce effort on those aspects of performance which are not rewarded. To test for this behavioral response to incentive schemes among the pilot physicians, we examined their performance during the time period of the study on two traditional quality measures. Of the 17 pilot physicians whose performance on mammogram and colorectal screening was reviewed in both January of 2001 and January of 2002, 10 of these physicians improved their screening rates and screening rates for the remaining seven were unchanged. Though far from conclusive, these data suggest that the pilot physicians did not reallocate effort away from preventive screening toward diabetes care.

Physician Feedback

In the spring of 2003, all physicians involved in this program were invited to a meeting to review results of this program and to elicit feedback on the initiative. The physicians raised a number of issues that offer insight on incentive program design, implementation, and the potential for health plan-physician collaboration.

Initially, the physicians reacted to the program with suspicion and distrust. They did not believe they would actually receive additional payments nor did they believe that the program would improve care. The physicians reported being surprised by their low scores on the baseline report and poor

compliance with the guidelines. In many cases, they also were confronted with poor documentation of clinical data in their records.

Most physicians responded to these discoveries by developing a clinical protocol checklist to improve documentation. When the physicians became aware that a number of patients on their registry were not compliant with office visits, they and their staff members made outreach calls to re-engage with these patients. The physicians reported that patients were pleasantly surprised and responsive to the outreach calls and additionally, that these interventions often prompted enrollment in disease management programs.

Two of the four physicians interviewed reported that they assigned patient monitoring and follow-up responsibilities to office staff and that these new responsibilities were well accepted by the staff. One of these physicians also shared their financial rewards with staff members. Finally, and perhaps most promising from a quality improvement perspective, were reports that physicians began to use the clinical protocol checklist for patients from other health plans even though financial rewards would not be forthcoming for these patients. This is a powerful indication that genuine improvement efforts were being integrated into the practice's workflow. Most physicians indicated they would use similar checklists for other diagnoses if they were available.

Physicians were asked to comment on the size of the potential reward. About half of the physicians felt that it was sufficient to cover the fixed costs of establishing or updating the infrastructure needed to deliver care according to the guideline but did not provide much remuneration beyond this. The other half of the group of pilot physicians felt that the size of the reward was substantial, had gotten their attention, and was influential. Privately, one physician remarked that the financial incentives were not large enough to compensate him for the opportunity cost of seeing more patients. This comment signals more than a lower bound estimate of how large incentives need to be to influence behavior; it demonstrates that some physicians are focused on practice finances and simply trying to cover their overhead to break even. We expect that physicians in these situations are unlikely to be influenced by uncertain rewards for quality improvement.

It is widely acknowledged that patient compliance with physician instructions is a very important determinant for performance on quality measures for chronic disease management. For example, a physician may urge her patient to obtain a diabetic retinal examination, but ultimately it is the patient who acts or does not act on this prescription. The issue of patient compliance becomes more complex when the rates of compliance are correlated with observable patient characteristics. One physician in this study

commented that low-income patients focused on meeting basic needs are unlikely to put a diabetic retinal exam at the top of their priority list. If the financial rewards to improved quality are large enough, then physicians might benefit financially from encouraging these challenging patients to leave their practice. One physician remarked that the outcome targets ($\text{HbA1c} < 7.5$) were unattainable for some of his diabetic patients. Note that a financial incentive system that only rewards the achievement of certain targets could aggravate this problem. More than one physician suggested that, because the health of diabetics is so strongly influenced by patient behaviors, the health plan should combine rewards to physicians with rewards to patients.

Physicians also remarked that their active participation in the incentive program might lead to uncomfortable moments with patients. Some physicians were uneasy with their patients knowing that they were receiving additional payments for what they were already supposed to be doing. Critics of pay for performance in health care have made this point as a general objection to paying for quality.

In one medical group, six out of eight physicians elected to participate in the incentive program. One of the six physicians remarked that participation by the majority of his colleagues encouraged him to participate and that the other two physicians cited increased paperwork as their primary reason for not participating. The belief that participation in the incentive program would unduly increase paperwork is both supported by and contradicted by physicians' reported experiences. Some physicians remarked that the documentation requirements of the program were not onerous and were easily integrated into existing or revised office procedures. Other physicians did not collect data in an ongoing fashion but concentrated this activity into one day in the office when they were not seeing patients. For the latter group of physicians, it is likely they perceived the administrative work associated with the program as a burden; these physicians also seemed less likely to delegate tasks to staff members and appeared less likely to integrate quality improvement activities into their daily routines.

DISCUSSION

Creating an infrastructure for office-based quality improvement requires that practice leaders recognize the importance of monitoring the health status of their patient population. Physicians are not easily persuaded that population

management is important until they are presented with data on their own practices and see for themselves the gaps between best practices and actual clinical care delivered in their office. This “epiphany” can trigger the creation and implementation of simple systems to measure, monitor, and track process and outcome measures of quality. These systems create new and expanded roles and responsibilities for staff, new procedures to follow-up on patient care, and interventions that take place outside of the 15–20 minutes office visit.

It is impossible to determine, from this study, the marginal effects of performance cash bonuses on physician behavior and the consequent changes in quality of care, separately from the other care management tools introduced (e.g., patient registries). It is highly unlikely that the payment alone would have stimulated the observed changes in performance on the composite score; however, it is also unlikely that as much focus and attention would have been brought to bear on diabetic care had no financial incentives been offered.

Initial results for this pilot project are both positive and encouraging. Yet, these very accomplishments raise important issues about scale and scope for quality incentive programs. At the current time, it is unknown whether paying for quality can or will lead to similar improvements in the care for patients with other chronic diseases and for patients with acute illnesses. We hypothesize that the likely success of quality incentive programs applied to other diseases will hinge on the availability, integrity, and legitimacy of performance measures.

The second thorny issue raised in this arena is how to transition from these initial successes in paying for quality to a fundamentally different model for provider and health plan reimbursement (Cutler 2003). The authors of the IOM report caution that although current payment methodologies may be made more supportive of quality improvement, fundamental change will be required to create significant and enduring incentives for quality improvement. System changes of this magnitude are always very difficult and complicated; however this particular change is likely to run afoul of the physicians’ demands that quality incentive payments be in addition to, and not in place of, current reimbursements for care delivery.

Third, recent research calls into question employers’ and consumers’ willingness to pay for quality—at least for clinical quality as it is defined in HEDIS. In the demonstration project reported on in this paper, IH experienced no net financial benefit from the quality improvements it instigated; in fact, in financial terms it lost money. In theory, quality improvement can lead to higher profits for health plans through two mechanisms: medical care cost savings or higher premiums. Existing research indicates chronic disease

management does not result in significant medical care cost savings until several years after program implementation (Beaulieu et al. 2003). Thus, research on the questions of whether and what types of quality consumers are willing to pay for is paramount for the widespread adoption of quality incentive programs.

Finally, the sample size of this pilot program was quite small and limits the potential generalizability of the results. It is important to note that physicians who participated in the pilot volunteered and thus may have been either more able or more motivated to improve. More research is needed to assess how the effects of financial rewards for quality vary with respect to factors such as patient population, health plan–physician contracts, physician practice characteristics, and incentive design.

CONCLUSIONS

The IH quality incentive pilot provides one blueprint for how to develop and implement an incentive program for chronic disease. This incentive program is unique in that it not only paid financial rewards for improvement, but also provided physicians with a support system for implementation. This infrastructure resulted in heightened awareness on the part of physicians of the gaps in their care processes and stimulated the adoption of a variety of practice level interventions to close the gaps. The incentive program, and the behaviors and investments it engendered, were associated with significant improvements in the quality of health care services delivered to diabetic patients and in physiological outcomes that are known to be predictive of improved long-term health status.

ACKNOWLEDGMENTS

The authors would like to acknowledge helpful comments and insights from two anonymous reviewers, Jean Manno, James Roistacher, Sharon Hewner, Dianne Hurren, Dr. William Major, and Meredith Rosenthal.

NOTE

1. Health Plan HEDIS scores for Comprehensive Diabetes Care were reported separately according to patient population: Medicare and commercial. These

separate scores were combined using the number of pilot study patients in each of these products as weights.

REFERENCES

- Beaulieu, N. D., D. Cutler, K. Ho, D. Horrigan, and G. Isham. 2003. *The Business Case for Diabetes Disease Management at Two Managed Care Organizations: A Case Study of HealthPartners and Independent Health Association (#610)*. New York: The Commonwealth Fund. Available at <http://www.cmf.org/>
- Bodenheimer, T., E. H. Wagner, and K. Grumbach. 2002a. "Improving Primary Care for Patients with Chronic Illness." *Journal of the American Medical Association* 288 (4): 1775–9.
- . 2002b. "Improving Primary Care for Patients with Chronic Illness: The Chronic Care Model, Part 2." *Journal of the American Medical Association* 288 (15): 1909–14.
- Cutler, D. M. 2003. *Your Money or Your Life: Strong Medicine for America's Health Care System*, chapter 9. New York: Oxford University Press.
- Institute of Medicine. 2001. "Aligning Payment Policies with Quality Improvement." In *Crossing the Quality Chasm*, chapter 8. Washington, DC: National Academy Press.
- Rosenthal, M., R. Fernandopulle, H. Ryu Song, and B. Landon. 2004. "Paying for Quality: Providers' Incentives for Quality Improvement." *Health Affairs* 23 (2): 127–41.