

# Effect of Physician-Specific Pay-for-Performance Incentives in a Large Group Practice

Sukyung Chung, PhD; Latha P. Palaniappan, MD, MS; Laurel M. Trujillo, MD; Haya R. Rubin, MD, PhD; and Harold S. Luft, PhD

In recent years, a growing number of health plans and payers in the United States have adopted pay-for-performance (P4P) mechanisms to encourage improvement in quality of care. Despite its promise and growing popularity, empirical evidence on the effectiveness of P4P is inconsistent. Some studies<sup>1-9</sup> have demonstrated the hoped-for positive effects, but others found no effect<sup>10-12</sup> or unintended negative effects.<sup>13</sup> These inconsistent findings may in part arise from heterogeneities in the scope of targeted quality dimensions and in the design and implementation of the program. Pay-for-performance programs are increasingly targeting multiple measures rather than focusing on 1 or 2 measures. Targeted measures are often process measures but sometimes include outcome measures.<sup>2,5,13</sup> Recent studies<sup>2,3,5-8</sup> tend to demonstrate small effects of P4P, but even these are not uniform across all measures evaluated.

Most P4P programs in published studies were designed and implemented by a payer rather than by the physicians subject to the incentives. The buy-in to such externally determined specific performance measures and associated incentives by physicians may be limited. Payer-driven incentive schemes are typically designed to work across a wide range of sites and data systems. This “least common denominator” approach may not always seem clinically applicable to those physicians who believe they have better data. Eligibility restrictions on incentive payments (eg, only applicable to enrollees of certain insurers) may lead to focused attention on just those patients or to a failure to change practice because too few patients would be involved. Studies<sup>1,6,9,11,12</sup> typically assess the effects of a modest group-level incentive compared with no incentives. Increasingly, studies<sup>2,3,5,7</sup> are assessing individual physician-level incentives again compared with no incentives. However, it is unknown whether the target of incentives (ie, group level vs physician specific) would make a difference. The present study explores this last question.

The primary objective of this study was to assess the effects of physician-specific P4P incentives versus group-level incentives on various quality measures. The P4P program examined in this study is similar to other P4P programs, but its specific implementation was physician led rather than payer driven. It was designed via a consensus process among representatives of the participating physicians with regard to the definitions of quality measures,

**Objective:** To assess the effect of a physician-specific pay-for-performance program on quality-of-care measures in a large group practice.

**Study Design:** In 2007, Palo Alto Medical Clinic, a multispecialty physician group practice, changed from group-focused to physician-specific pay-for-performance incentives. Primary care physicians received incentive payments based on their quarterly assessed performance.

**Methods:** We examined 9 reported and incentivized clinical outcome and process measures. Five reported and nonincentivized measures were used for comparison purposes. The quality score of each physician for each measure was the main dependent variable and was calculated as follows: Quality Score = (Patients Meeting Target / Eligible Patients) × 100. Differences in scores between 2006 and 2007 were compared with differences in scores between 2005 and 2006. We also compared the performance of Palo Alto Medical Clinic with that of 2 other affiliated physician groups implementing group-level incentives.

**Results:** Eight of 9 reported and incentivized measures showed significant improvement in 2007 compared with 2006. Three measures showed an improvement trend significantly better than the previous year’s trend. A similar improvement trend was observed in 1 related measure that was reported but was nonincentivized. However, the improvement trend of Palo Alto Medical Clinic was not consistently different from that of the other 2 physician groups.

**Conclusions:** Small financial incentives (maximum, \$5000/year) based on individual physicians’ performance may have led to continued or enhanced improvement in well-established ambulatory care measures. Compared with other quality improvement programs having alternative foci for incentives (eg, increasing support for staff hours), the effect of physician-specific incentives was not evident.

(*Am J Manag Care.* 2010;16(2):e35-e42)

**In this article**  
Take-Away Points / e36  
**Published as a Web Exclusive**  
www.ajmc.com

designed via a consensus process among representatives of the participating physicians with regard to the definitions of quality measures,

**For author information and disclosures, see end of text.**

### Take-Away Points

Various well-established quality-of-care indicators demonstrated continued improvement during a trial of physician-specific quality financial incentives.

- In the context of other organization-level quality improvement efforts, physician-specific incentives seem to have some incremental effect of improving quality of care.
- These changes occurred in a practice setting relying on fee-for-service payment to physicians and with many of the quality goals unassociated with additional tests or procedures.
- Alternative foci of incentives for quality improvement (eg, increasing staff hours to assist physicians' quality improvement activities) need to be explored.

### Design of the Physician Incentive Program

In 2007, all PCPs ( $n = 179$ ) at PAMC practicing family medicine, general internal medicine, or pediatrics participated in the physician incentive program. The bonus amount was based on individual physicians' performance on 15 ambulatory quality measures, with a composite score calculated using an algorithm developed by the incentive program leadership.<sup>15</sup> In brief, the physicians set targets for each measure. Physicians received varying points for achieving minimal, average, and stretch goals based on the percentage of their patients achieving the target. The bonus was based on the percentage of potentially achievable points actually earned. The maximum achievable bonus was \$5000/year, or about 2% of the PCP annual salary. The design and implementation of the program were discussed at the physician partnership meeting and at primary care department head meetings.

Since 2003, PAMF has been participating in the P4P program sponsored by several California health plans and operated by the Integrated Healthcare Association (IHA) (<http://www.ih.org>). PAMF retained a portion of the IHA P4P incentive for its organization-wide quality improvement (QI) efforts, and the remainder (roughly \$4000 per physician) was distributed to each of 3 physician groups. The PAMF portion went to support the considerable central organizational services provided by PAMF for QI interventions on behalf of physicians.

Three physician groups independently decided on the allocation of the remainder of the funds. Until 2006, the allocations to the physician groups were not passed on to individual physicians. In 2007, PAMC decided to distribute a portion of the IHA payments to individual PCPs based on their performance scores on measures that were internally defined by the group and had been used at least since 2005. The other 2 physician groups continued using the IHA measures and definition of eligible patient populations in assessing group-level or department-level performance.

### Quality Monitoring and Reporting

Although the physician-specific financial incentive was newly implemented in 2007, monitoring and quarterly reporting of quality indicators in the PAMC system had been in place since 2003. Physicians were alerted by e-mail with an electronic link to a detailed quality score workbook with their scores, peer physicians' scores, and rank relative to other physicians in a distribution curve for each quality measure. In these reports, physician identities are disclosed to each other.

the inclusion and exclusion rules among patients for each measure, and the incentive formulas. Although the measures implemented are similar to those required by and reported to payers, the physicians used criteria for inclusion and exclusion and for targets of success that they believed were more relevant to their clinical practice. For example, completion of colon cancer screening in their version could be confirmed by patient self-report or by documents verifying completion in a different institution rather than just through services provided by the group. For patients with diabetes mellitus, the physicians chose a more stringent glycosylated hemoglobin target of 7.0% instead of the externally set target of 8.0%. Most important, all patients of the primary care physicians (PCPs) (regardless of insurer) were considered for the performance evaluation so that work flow could be altered for all patients. The data used in the study came from electronic health records, which generally provide more accurate and precise data on clinical procedures and outcomes than billing data.<sup>14</sup>

## METHODS

### Study Setting

The study was conducted at Palo Alto Medical Foundation (PAMF), a not-for-profit healthcare organization. In 2007, PAMF contracted with the following 3 multispecialty physician groups: Palo Alto Medical Clinic (PAMC), Camino Medical Group, and Santa Cruz Medical Group. Although 25% to 30% of PAMF patients are enrolled in capitated programs, the physicians in the groups contracting with PAMF were paid based on relative value units of service regardless of the patient's coverage. All 3 physician groups, located at clinics in adjacent counties, had a roughly similar mix of primary care and specialty physicians and served patients of similar demographic composition. The physician-specific P4P incentive program was implemented at PAMC in 2007. PAMC provides coverage at 5 clinics operating within 3 counties in the San Francisco Bay Area, California, serving approximately 13% of the general population in the underlying geographic area, with low patient turnover (3% per year).

## Effect of Physician-Specific Pay-for-Performance Incentives

Before the implementation of the physician incentive program at PAMC, the program leadership convened several meetings to decide on the performance measures to use, details about each measure, and target levels of performance to incentivize and then selected 15 clinical measures representing clinical outcomes and processes. The present study focuses on 9 reported and incentivized clinical outcome and process measures that had been already evaluated and reported at least since 2005. The other 6 measures, which were specific to pediatric patients, were newly adopted for the 2007 program; the present study excludes these 6 measures without the prior year data needed for identification of the program effect. The 9 incentivized measures analyzed in this study were 3 outcome measures for patients with diabetes and 6 process measures (Table 1).

Several of 9 measures were similar to those used in the IHA's P4P program (<http://www.iha.org>), which was developed from evidence-based practice guidelines such as the 2007 Healthcare Effectiveness and Data Information Set indicators by the National Committee for Quality Assurance. The specific definitions of the measures used in the PAMC program were somewhat different, reflecting the group's organizational goals, high standard of quality of care, and information technology capacity. For example, it set stricter thresholds for the diabetic control indicators than did the IHA (eg, 100 vs <130 mg/dL for low-density lipoprotein cholesterol [LDL-C] control; to convert cholesterol level to millimoles per liter, multiply by 0.0259). In contrast, patients whose completed screening tests could be verified through other providers were coded as completed, whereas the IHA required that the tests be performed by the physician group.

Our analysis uses 5 comparison and nonincentivized measures reflecting similar dimensions of care reported since 2005 or earlier but not included in the physician-specific incentive program. These include the following 5 measures: 4 outcome measures (1 for patients with hypertension and 3 for patients with diabetes [the same indicators as were incentivized but with less stringent targets]) and 1 process measure focused on a different target population (patients with hypertension) (Table 1).

The quality score of each physician was calculated for each measure with denominators of 6 or more eligible patients in a physician's panel as follows: Quality Score = (Numerator/Denominator) × 100, where the numerator is the number of patients who received the recommended care, and the denominator is the number of patients who were eligible for the recommended care (definitions for each measure are given in Table 1). The purpose of the physicians setting the minimum criterion for the denominator in the incentive design was to prevent the percentage score from being dominated by a few

cases; we follow their approach. With the help of a health maintenance compliance reminder system embedded in the electronic health record and diabetes and asthma registry data, physicians could easily identify their eligible patients for each measure.

### Statistical Analysis

The unit of analysis in the primary analysis is the physician, eliminating concerns about a changing mix of providers over time. We first analyzed time trends in quality scores over 3 years (2005-2007) for each measure. We did not have access to patient-level data, so it was impossible to control for within-category comorbidities. However, the high stability of patient panels suggests that comorbidity patterns in a physician's panel do not vary markedly over time and that the same patients are included for multiple years for certain measures. To take into account within-physician correlations across quality measures, we estimated physician-level random-effects models. Statistical significance was set at  $P < .05$ .

To assess the effects of temporal trends, we descriptively compared the trend in scores among the PAMC physicians relative to those of the other 2 physician groups that had not implemented a physician-specific incentive program. For this purpose, we used the mean performance scores for the following 3 measures in the IHA P4P program that were defined identically among the 3 physician groups for 3 years (2005-2007): asthma controller prescribing, cervical cancer screening, and *Chlamydia* screening. The IHA scores are not available at a physician-specific level and may reflect some changes in providers across years. Because the IHA data are aggregated, we cannot perform the same types of statistical tests as were performed for the PAMC.

## RESULTS

All PCPs practicing at 5 PAMC locations were eligible for the program. Among 179 physicians, 167 were included in the study; 12 had insufficient qualifying patients for various reasons (eg, medical leave or too few patients). Physicians were more likely to be in family medicine (42%) or general internal medicine (34%) than in pediatrics (24%). Most physicians (152 in 2005 and 169 in 2006) also had data for the equivalent measures in the previous years; 148 had data for all 3 years.

Eight of 9 reported and incentivized measures showed significant improvement in 2007 compared with 2006. The rate of improvement was significantly better than the previous year's trend (2006 compared with 2005) for the following 3 measures: blood pressure control for patients with diabetes, colon cancer screening, and documentation of tobacco use

■ **Table 1.** Definition of Quality Measures

Quality Measure	Numerator	Denominator <sup>a</sup>
<b>Included in the physician-specific incentive program</b>		
<b>Clinical outcome</b>		
Diabetes mellitus A1C control (<7%)	Most recent A1C is <7%	Diabetes roster patients whose A1C was measured within the past 6 mo <sup>b</sup>
Diabetes blood pressure control (≤130/80 mm Hg)	Mean (of ≤3 most recent) blood pressures at clinic visits for the past 12 mo is ≤130/80 mm Hg	Diabetes roster patients whose blood pressure was measured within the past 12 mo <sup>b</sup>
Diabetes LDL-C control (<100 mg/dL)	Most recent LDL-C is <100 mg/dL	Diabetes roster patients whose LDL-C was measured within the past 12 mo <sup>b</sup>
<b>Clinical process</b>		
Asthma controller prescribing	Long-term controller (corticosteroid inhaler, cromolyn sodium, xanthine, or leukotriene modifier) is currently prescribed	Patients in asthma registry <sup>c</sup>
Cervical cancer screening	Pap smear performed within the past 3 y or hysterectomy ever performed <sup>d</sup>	Women 21-64 y who had ≥1 visit in the past 24 mo
<i>Chlamydia</i> screening	<i>Chlamydia</i> screening performed within the past 12 mo <sup>d</sup>	Women 16-25 y who had ≥1 visit in the past 24 mo and are reported as sexually active <sup>e</sup>
Colon cancer screening	Fully screened for colon cancer by colonoscopy within the past 10 y or by both flexible sigmoidoscopy and stool occult blood within the past 12 mo <sup>d</sup>	Patients 51-75 y who had ≥1 visit in the past 24 mo
Documentation of tobacco use history	Patients who have something other than “not asked” entered for tobacco history or who have text entered in “tobacco comment”	Patients ≥12 y who had ≥2 visits with the PCP in the past 24 mo, with ≥1 visit in the last quarter
Measurement of height and weight	Height and weight were measured and recorded since the patient’s 18th birthday	Patients ≥18 y who had ≥1 visit in the past 24 mo
<b>Reported to the physicians but not included in the physician-specific incentive program</b>		
<b>Clinical outcome</b>		
Diabetes A1C control (<8%)	Most recent A1C is <8%	Diabetes roster patients whose A1C was measured and recorded within the past 6 mo <sup>b</sup>
Diabetes blood pressure control (≤140/90 mm Hg)	Mean (of ≤3 most recent) blood pressures at clinic visits for the past 12 mo is ≤140/90 mm Hg	Diabetes roster patients whose blood pressure was measured and recorded within the past 12 mo <sup>b</sup>
Diabetes LDL-C control (<130 mg/dL)	Most recent LDL-C is <130 mg/dL	Diabetes roster patients whose LDL-C was measured and recorded within the past 12 mo <sup>b</sup>
Hypertension blood pressure control (≤140/90 mm Hg)	Mean (of ≤3 most recent) blood pressures at clinic visits within the past 12 mo is ≤140/90 mm Hg	Hypertension roster patients whose blood pressure was checked and recorded within the past 12 mo <sup>b</sup>
<b>Clinical process</b>		
Hypertension blood pressure check	Blood pressure was measured and recorded within the past 12 mo	Hypertension roster patients <sup>b</sup>

A1C indicates glycosylated hemoglobin; LDL-C, low-density lipoprotein cholesterol; PCP, primary care physician.  
 SI conversion factor: To convert cholesterol level to millimoles per liter, multiply by 0.0259.  
<sup>a</sup>All patients in the denominator had an active PCP listed.  
<sup>b</sup>Diabetes and hypertension roster patients are 18 years or older who have arrived for at least 1 visit in the past 24 months, have an active PCP listed, and have a diabetes or hypertension diagnosis on the active problem list.  
<sup>c</sup>Patients in asthma registry are aged 0 to 56 years who have arrived for at least 1 visit in the past 24 months, have an active PCP listed, and have a persistent asthma diagnosis or 4 visits in the past 2 years with an asthma diagnosis on the active problem list.  
<sup>d</sup>For cervical cancer, *Chlamydia* screening, and colon cancer screening, 3 types of health maintenance overrides are also counted as completed (done, external completion–patient reported, or external completion–confirmed). Overrides allow physician input on patient completion of testing outside of the Palo Alto Medical Foundation system.  
<sup>e</sup>Sexually active if meet any of the following 4 criteria: (1) sexual history in the electronic health record, (2) a procedure code within the past 12 months indicating sexual activity (pregnancy, miscarriage, abortion, sexually transmitted disease screening or treatment, infertility, or Pap smear), (3) a diagnosis code within the past 12 months indicating sexual activity, or (4) a prescription within the past 12 months for birth control pills or diaphragm.

## Effect of Physician-Specific Pay-for-Performance Incentives

**Table 2.** Trend in Quality Scores Before and After Physician Incentive Program Implementation

Quality Measure	Quality Score, %			Difference in Quality Score <sup>a</sup>		Difference in Difference in Quality Score <sup>b</sup>
	2005 (n = 152)	2006 (n = 169)	2007 (n = 167)	2006–2005 = 0	2007–2006 = 0	(2007–2006)–(2006–2005) = 0
<b>Physician-specific incentive program</b>						
Diabetes mellitus A1C control (<7%)	58	60	62		c	
Diabetes blood pressure control (≤130/80 mm Hg)	47	49	53	c	c	c
Diabetes LDL-C control (<100 mg/dL)	60	63	60	c	c	(c)
Asthma controller prescribing <sup>a</sup>	91	92	92		d	
Cervical cancer screening <sup>a</sup>	75	77	79	c	c	
<i>Chlamydia</i> screening <sup>a</sup>	36	37	38		d	
Colon cancer screening	38	40	47	c	c	c
Documentation of tobacco use history <sup>a</sup>	72	75	79	c	c	c
Measurement of height and weight	68	70	73	c	c	
<b>Report to the physicians but not included in the physician-specific incentive program</b>						
Diabetes A1C control (<8%)	81	81	83	d	c	
Diabetes blood pressure control (≤140/90 mm Hg)	77	78	81	c	c	
Diabetes LDL-C control (<130 mg/dL)	86	88	87	c		(c)
Hypertension blood pressure control (≤140/90 mm Hg)	64	67	72	c	c	c
Hypertension blood pressure check	90	90	90	(d)		

A1C indicates glycosylated hemoglobin; LDL-C, low-density lipoprotein cholesterol.  
 SI conversion factor: To convert cholesterol level to millimoles per liter, multiply by 0.0259.  
<sup>a</sup>Statistics based on the results from the multilevel mixed-effects linear regression (z statistics); parentheses are used when the difference is negative.  
<sup>b</sup>Statistics based on F test of equivalence of coefficients; parentheses are used when the difference is negative.  
<sup>c</sup>Significant at 1%.  
<sup>d</sup>Significant at 5%.

history (Table 2). For LDL-C level control among patients with diabetes, the change in 2006-2007 reversed the improvement seen in 2005-2006. This reverse trend at the end of 2006 was apparently due to a change in the laboratory cholesterol analyzer, which produced systematically lower LDL-C values. Accelerating improvement was observed in 1 of 5 nonincentivized measures examined, namely, blood pressure control for patients with hypertension.

Trends in the 3 IHA P4P measures assessed consistently across the 3 physician groups did not clearly demonstrate any greater improvement in 2007 at the PAMC compared with 2006 or compared with trends in the other 2 physician groups. These results are shown in the Figure.

### DISCUSSION

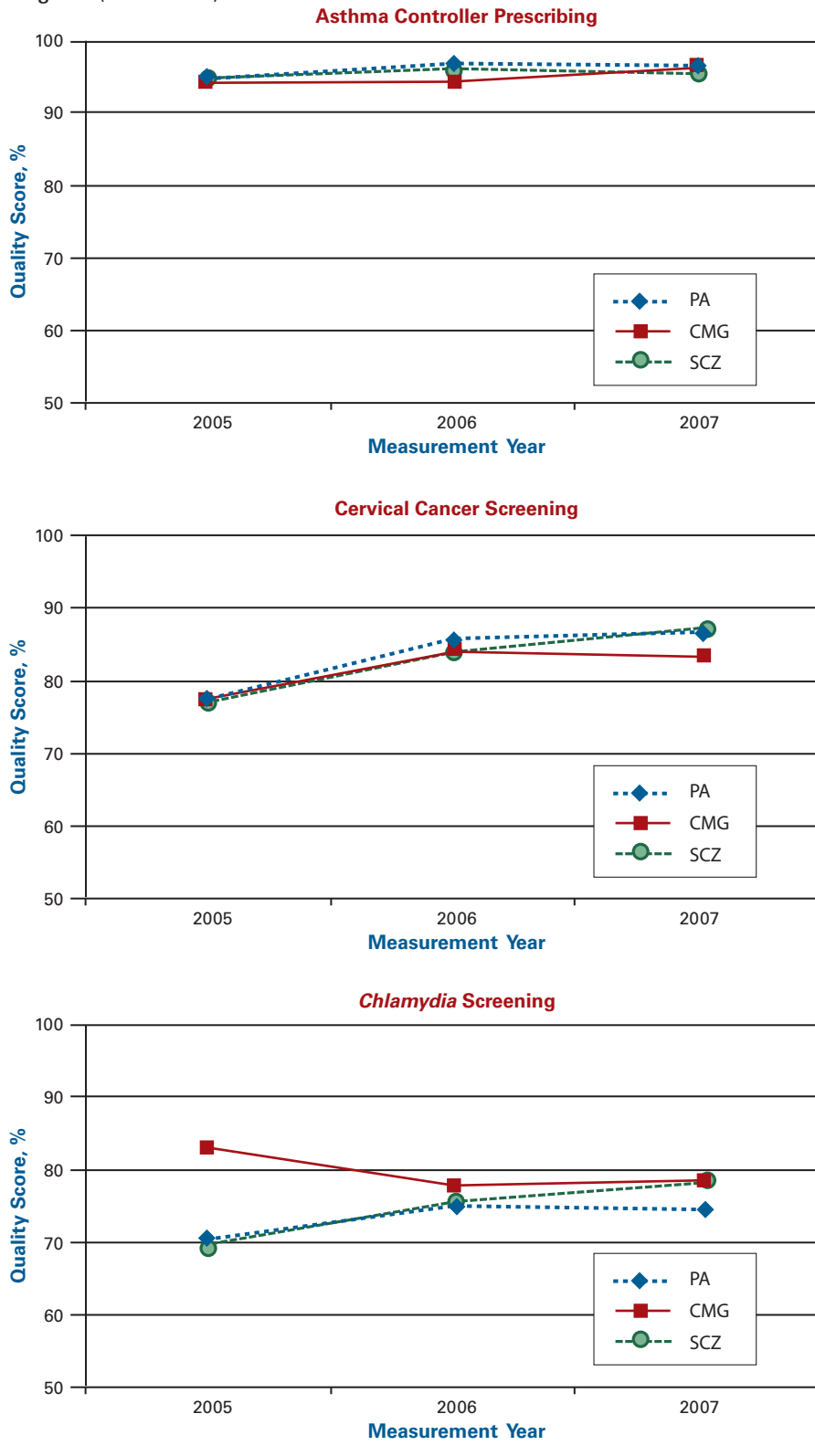
With an individual physician-level P4P program, we observed improvement in some reported and incentivized

quality measures during the intervention period and in comparison with the previous years' trend. However, we observed no difference in consistently measured quality indicators between PAMC and the other 2 physician groups without physician-specific incentives. However, all 3 physician groups implemented various other QI activities, and it is impossible to rule out other extraneous but site-specific factors that may have influenced these results. Nonetheless, these findings suggest that physician-specific financial incentives, among other strategies, could be effective to improve quality of care.

Our study focused on a program designed and implemented with a high level of participating physicians' engagement and is notable for the inclusion of all patients and for a more comprehensive scope of measures than most P4P programs. The quality reporting system allowed the physicians to easily track their eligible patients and peer physicians' performance. Given this context, the physician group may have



**Figure.** Comparison With Other Physician Groups' Quality Scores for Similar Measures Used in the Integrated Healthcare Association Pay-for-Performance Program (2005-2007)



PA indicates Palo Alto Medical Clinic; CMG, Camino Medical Group; SCZ, Santa Cruz Medical Group.

expected positive outcomes. However, because various QI programs had already been implemented and because organization-level incentives had been in place at the study site for many years, it may have been unrealistic to expect substantial additional improvement.

The relative similarity in trends between PAMC and the other 2 physician groups may be happenstance. The other physician groups, with no internally defined P4P program and with less comprehensive data, could only focus on those patients identified by the IHA collaborating health plans, roughly 30% of the total. To the extent that some QI efforts may have involved extra effort in outreach to patients or in patient education, this means that the same incremental resources could be focused on fewer patients, leading to the expectation of a greater effect outside of PAMC.

Our findings are consistent with some recently published P4P studies. Similar to our finding of accelerated improvement in 2007 for blood pressure control among patients with diabetes, Beaulieu and Horrigan<sup>2</sup> reported positive glycosylated hemoglobin and LDL-C outcomes among this patient population. Other studies found similar results regarding improved cancer screening<sup>5</sup> and found no improvement in asthma controller prescribing.<sup>7</sup> Another study<sup>9</sup> examining group-level incentives echoes our finding of improved documentation

## Effect of Physician-Specific Pay-for-Performance Incentives

of tobacco use history. Taken together, well-designed P4P incentive programs seem to have modest positive effects on improving targeted dimensions of quality of care.<sup>16</sup>

The estimated improvement due to the physician-specific incentives in our study may be conservative for at least 2 technical reasons. First, for the percentage score in a particular quarter, the denominator was the number of patients ever seen by a PAMC physician for periods ranging from the previous 6 months to the past several years, depending on the measure. If only those patients who were newly qualified to be in the denominator were included, the effect would have been substantially larger. The lack of variation due to the denominator effect might also have lessened physicians' response to the incentive program. Second, by design, physicians already had been receiving bonuses provided to the entire group based on similar performance measures. We might have found a larger effect of physician-specific incentives if the physician group had had no financial incentives at all before the provision of physician-specific incentives. Other studies<sup>2,3,5,7</sup> typically used such comparisons.

Although we found an improvement in quality among some measures with a small bonus, the same or better improvement might have been achieved through other uses of the funds (eg, increasing staff hours and investment in additional information technology to easily track target patients). Given limited physician office hours and the increasing number and complexity of clinical guidelines and recommendations, provision of extra resources to substitute or complement physicians' time may work better than modest monetary incentives given solely to physicians. The lack of observable difference in improvement relative to the other physician groups adopting different QI strategies also suggests potential advantages of other forms of incentives. Quality improvement may well depend on a wide range of interventions, including data infrastructure, physician and staff engagement, changes in work flow, and incentives. Future studies should investigate the effectiveness of alternative approaches to achieve QI.

Several limitations of our study merit discussion. First, we did not have a contemporaneous comparison group at the same study site receiving only performance reporting or group-level incentives; such a study design was unacceptable by the participating physicians. Instead, we descriptively compared group-level aggregate performance scores with scores of other physician groups serving similar patient populations. This limits the inferences we can make about the direct effect of the program. Second, experiences from one medical group cannot be generalized to those of other groups with different sizes, locations, information technology capacities, and "culture." PAMC had been already providing high-quality services with the help of advanced information technology

(eg, the electronic health record database, diabetes registry, asthma registry, and reporting system of comprehensive quality measures); the implementation and the effects of a similar physician-specific incentive program in other groups with different capacity might be different from what is reported herein. Third, the incentive program was discontinued after 1 year due to the merger of the 3 physician groups and the complexity of developing a uniform compensation scheme. This decision was made long before any results of the incentive program became available. Therefore, we are unable to assess how the incentive effect would change over time.

In future studies, important variations in payment schemes such as varying the amount and risk (ie, bonus vs penalty) of payment should be examined. The types of programs that may work well in a highly integrated medical group may be different from those that are best for independently operating practices. Further investigation about what is the driving force of the improvement or lack of improvement across the various incentivized measures is also needed. Specific physician and group characteristics related to responsiveness to P4P, as well as within-physician correlation across measures, should be understood to better tailor a program to varying practice settings.

In conclusion, physician-specific financial incentives (albeit small) for higher-quality care applied to well-established ambulatory care measures and implemented in the context of other ongoing organization-level QI efforts were associated with incremental improvements in aspects of quality of care that were incentivized. The effects of broader-scope performance-based physician compensation schemes and other types of incentives for QI (eg, increasing staff hours to assist physicians' QI activities) need to be explored and compared.

### Acknowledgments

Earlier versions of this study were presented at the Second Agency for Healthcare Research and Quality Annual Conference, Washington, DC, September 7-10, 2005, the Fourth National Pay for Performance Summit, San Francisco, CA, March 9-11, 2009, the 137th Annual Meeting of the American Public Health Association, 2008 Annual Meeting and Exposition, San Diego, CA, October 25-29, 2008, and the 26th AcademyHealth Annual Research Meeting, Chicago, IL, June 28-30, 2009. We thank the participants of the meetings for their comments.

**Author Affiliations:** From the Research Institute, Palo Alto Medical Foundation (SC, LPP, LMT, HSL), Palo Alto, CA; Research and Evaluation (HRR), Palo Alto, CA; and School of Medicine, Johns Hopkins University (HRR), Baltimore, MD.

**Funding Source:** This study is derived from work supported under contract HHS290200600023I from the Agency for Healthcare Research and Quality.

**Author Disclosure:** The authors (SC, LPP, LMT, HRR, HSL) 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 (SC, LPP, LMT, HRR, HSL); acquisition of data (SC, LPP, HRR, HSL); analysis and interpretation of data (SC, LPP, LMT, HRR); drafting of the manuscript (SC, LPP, LMT, HSL); critical revision of the manuscript for important intellectual content (SC, LPP, HRR, HSL); statistical analysis (SC, HRR); provision of study materials or patients (LPP); obtaining funding (LPP, HRR, HSL); administra-

tive, technical, or logistic support (LPP, HRR); and supervision (LPP, HRR, HSL).

**Address correspondence to:** Sukyung Chung, PhD, Research Institute, Palo Alto Medical Foundation, 795 El Camino Real, Palo Alto, CA 94301. E-mail: chungs@pamfri.org.

## REFERENCES

1. Amundson G, Solberg LI, Reed M, Martini EM, Carlson R. Paying for quality improvement: compliance with tobacco cessation guidelines. *Jt Comm J Qual Saf.* 2003;29(2):59-65.
2. Beaulieu ND, Horrigan DR. Putting smart money to work for quality improvement. *Health Serv Res.* 2005;40(5, pt 1):1318-1334.
3. Doran T, Fullwood C, Gravelle H, et al. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med.* 2006;355(4):375-384.
4. Fairbrother G, Siegel MJ, Friedman S, Kory PD, Butts GC. Impact of financial incentives on documented immunization rates in the inner city: results of a randomized controlled trial. *Ambul Pediatr.* 2001;1(4):206-212.
5. Gilmore AS, Zhao Y, Kang N, et al. Patient outcomes and evidence-based medicine in a preferred provider organization setting: a six-year evaluation of a physician pay-for-performance program. *Health Serv Res.* 2007;42(6, pt 1):2140-2159.
6. Kouides RW, Bennett NM, Lewis B, Cappuccio JD, Barker WH, LaForce FM; Primary-Care Physicians of Monroe County. Performance-based physician reimbursement and influenza immunization rates in the elderly. *Am J Prev Med.* 1998;14(2):89-95.
7. Levin-Scherz J, DeVita N, Timbie J. Impact of pay-for-performance contracts and network registry on diabetes and asthma HEDIS measures in an integrated delivery network. *Med Care Res Rev.* 2006;63(1 suppl):14S-28S.
8. Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. *JAMA.* 2005;294(14):1788-1793.
9. Roski J, Jeddelloh R, An L, et al. The impact of financial incentives and a patient registry on preventive care quality: increasing provider adherence to evidence-based smoking cessation practice guidelines. *Prev Med.* 2003;36(3):291-299.
10. Fairbrother G, Hanson KL, Friedman S, Butts GC. The impact of physician bonuses, enhanced fees, and feedback on childhood immunization coverage rates. *Am J Public Health.* 1999;89(2):171-175.
11. Hillman AL, Ripley K, Goldfarb N, Nuamah I, Weiner J, Lusk E. Physician financial incentives and feedback: failure to increase cancer screening in Medicaid managed care. *Am J Public Health.* 1998;88(11):1699-1701.
12. Hillman AL, Ripley K, Goldfarb N, Weiner J, Nuamah I, Lusk E. The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care. *Pediatrics.* 1999;104(4, pt 1):931-935.
13. Shen Y. Selection incentives in a performance-based contracting system. *Health Serv Res.* 2003;38(2):535-552.
14. Tang PC, Ralston M, Arrigotti MF, Qureshi L, Graham J. Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *J Am Med Inform Assoc.* 2007;14(1):10-15.
15. Chung S, Palaniappan LP, Wong E, Rubin HR, Luft HS. Does the frequency of pay-for-performance payment matter? Experience from a randomized trial. *Health Serv Res.* 2009 Dec 31 [Epub ahead of print].
16. Rosenthal MB. P4P: rumors of its demise may be exaggerated. *Am J Manag Care.* 2007;13(5):238-239. ■