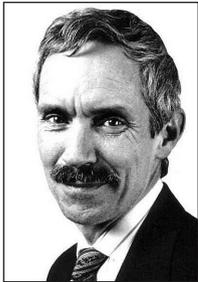




Pay for Performance: Two Critical Steps Needed to Achieve a Successful Program



L. Gregory Pawlson, MD, MPH
*Executive Vice President
NCQA
Washington, DC*

One year ago in *Decision Maker News in Managed Care*, an article entitled “Measuring Physician Performance, Enhancing the Quality of Care, and Reducing Costs” reported on the results of a group discussion at the Healthcare Summit in February 2003. This article builds on the framework of that report with a focus on 2 practical and critical steps essential to building a pay-for-performance (P4P) program that will actually impact quality and cost.

Over the past 18 months, there has been a further acceleration in effort by both purchasers and plans to create incentives for providers (hospitals and physicians) that reward quality and efficient use of resources. At this time, there are about 40 P4P programs being implemented by health plans and employer coalitions.¹ By 2006, at least 80 health plans covering 60 million patients are expected to offer such programs.²

Our current payment system clearly rewards providers for seeing patients and doing procedures, something they have been very successful at with the result that costs have increased. Paying providers differentially for enhanced quality and efficient use of resources appears to be a major step in creating market forces that can fundamentally affect quality and cost. This is in stark contrast to efforts that are unlikely to have any long-term effects on expenses, such as

shifting costs to patients via higher deductibles (ignoring the fact that most costs are concentrated on patients who spend more than \$2000/year on healthcare) or that use steep provider discounts, which frequently just shifts costs to insurers who have smaller market shares or less clout. However, successfully impacting long-established and ingrained behaviors, even with monetary incentives, is not easy. While the conceptual framework of P4P programs is relatively simple, a program that will actually impact both quality and cost in a positive way is much more complicated.

PAY FOR PERFORMANCE PROGRAM

Although a number of components are needed in P4P programs to actually change provider behavior, such as setting rewards to a level that gets provider attention, involving providers in the process but not giving them veto power, and staying the course after the initial resistance appears, there are 2 components that are critical. They are (1) selecting and using a standardized set of the “right” measures, and (2) collaborating with other plans and purchasers within a given region or market.

Critical steps. In terms of selecting the “right” measures, it should be remembered that the main purpose of P4P programs is to affect the providers’ behavior. Paying out bonuses to higher performing providers is alone not the answer. It is critical that both the measures and the way they are collected and analyzed are reli-

able, valid, understandable, and improvable at the level of the clinician or hospital. For example, if the sample size is too small (there is growing empiric evidence that 25-30 is adequate for most purposes) or if the measures are not easily understood by the provider, no matter how many vivid colors are used to display the results, they will end up in the waste basket.

Health plan claims data. When using health plan claims data, there is also the issue of the accuracy (or perhaps more precisely the clinicians’ believability) of the data source. The problem with “top down” data use has been ameliorated in some instances by expanding the traditional visit and procedure claims data to include laboratory and pharmacy claims and laboratory results. Most of the current P4P programs use Health Employer Data and Information Set clinical measures,³ along with the physician portion of the Consumer Assessment of Health Plans,⁴ or the Ambulatory Care Evaluation Survey,³ and some measure documentation of the use of information technology (IT) and/or other systems using the Physician Practice Connection (PPC)⁵ or a similar tool (Figure). A few P4P programs are using measures of efficiency or resource use, most often employing a proprietary grouping or risk adjustment methodology.

By contrast, projects such as Bridges to Excellence (BTE), which is an initiative of large employers who are working in conjunction with Aetna, Massachusetts Blue Cross Blue Shield, Tufts Health Plan, United Healthplan, Harvard Pilgrim Healthplan, and Humana, use a “bottom up” approach to drive

FIGURE⁶

Pay for Performance Sample Weights and Benchmarks

The Integrated Healthcare Association–led program bases its bonuses on physicians reaching clinical benchmarks, achieving technology adoption, and patient experience. Listed below are some of the benchmarks that are used to calculate the bonuses and the percentages given for each area.

Clinical metrics

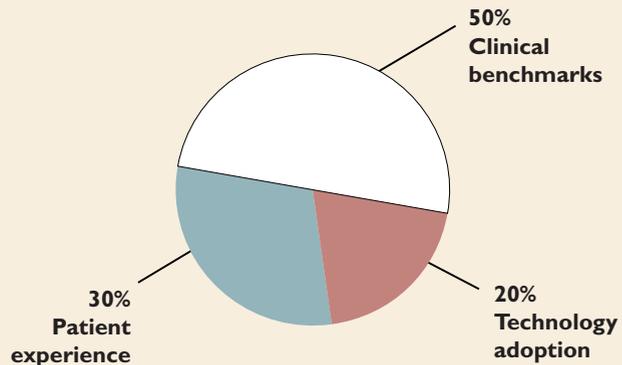
- Cancer screenings
- Childhood immunizations
- Disease management
- Diabetic HbA_{1c} screenings

Technology metrics

- Ability to integrate clinical data
- Adoption of computerized physician order entry
- Use of decision-supported software

Patient experience metrics

- Member satisfaction surveys



Adapted from: PriceWaterhouseCoopers. 2004.

improvement in 3 areas: diabetes care, cardiovascular care, and clinical patient care management systems.² They are using data reported by individual physicians or physician groups (with random audits to verify the information) extracted from external reviews of paper medical records as alternatives or even replacements for plan-level claims data. BTE uses the American Diabetes Association–National Commission for Quality Assurance (NCQA) Diabetes Recognition Program,⁷ the American Health Association–American Stroke Association–NCQA Heart Stroke Recognition Program⁸ (Table), and the PPC tool, which assesses the presence and functionality of office practice systems (IT, reminders, practice design, etc). The advantage of this approach is that it avoids physician resistance to use of administrative measures, and posts the review and reporting of data back to the physician practice level.

Cooperation between plans. The second critical factor is the need for

cooperation between plans, or between those organizations that can collect and provide standardized measurement of quality at the physician level. Both anecdotal and empiric data⁹ indicate the most common response from physicians getting multiple ratings or feedback from many health plans or other sources is to simply ignore the feedback. While these same physicians may be pleased to receive bonus payments from these plans, the likelihood that they will act on the information is small. As noted previously, P4P programs that do not drive improvement among providers are a waste of time and effort.

One approach to collaboration is through projects that create pooling of data across plans and/or providers within a given market. This effort can be led by coalitions of purchasers, such as the recently announced “Consumer-Focused Purchasing” program (a purchaser-led P4P program that is slated to be launched in multiple regional markets and which was spearheaded by Arnold Milstein, MD, and col-

leagues at Mercer)¹⁰; by providers (the Integrated Healthcare Association is leading a project in California that involves 6 health plans [Aetna, Blue Cross of California, Blue Shield of California, Cigna, Health Net, and PacifiCare] and more than 200 medical groups)¹¹; or through independent organizations like the Massachusetts Health Quality Partnership.¹² While remaining data differences between plans such as having unique provider identification numbers and variations in certain data elements make this problematic in some locations, it clearly provides a path past the barrier of small numbers or fragmented data feedback. Moreover, these projects, while more complex than “going it alone” often bring together the key stakeholders in an activity that can result in wins for purchasers, plans, providers—and ultimately, patients.

While antitrust and other legal concerns must be explored, the results from the Integrated Healthcare Association’s P4P program in California would suggest that if

TABLE⁸

Heart/Stroke Recognition Measures for Secondary Prevention of Heart Attack or Stroke

Standards	Criteria	Points
Blood pressure control (<140/90 mm Hg)	75% of patients in sample	10.0
Blood pressure results	Credit toward numerator*	
<140/90 mm Hg	1.00	
<145/90 or <140/95 mm Hg	0.75	
<145/95 mm Hg	0.50	
≥145/95 mm Hg	0.00	
Complete lipid profile	80% of patients in sample	10.0
Cholesterol control (<100 mg/dL)	50% of patients in sample	10.0
Low-density lipoprotein results	Credit toward numerator*	
<100 mg/dL	1.00	
100-109 mg/dL	0.75	
110-119 mg/dL	0.50	
120-129 mg/dL	0.25	
≥130 mg/dL	0.00	
Use of aspirin or other antithrombotic	80% of patients in sample	10.0
Smoking status and cessation advice or treatment	80% of patients in sample	10.0
	Total points:	50.0
Points needed to achieve recognition		40.0

*Denominator = number of charts reviewed in sample.

Adapted with permission from AHA/ASA Heart Stroke Recognition Program.

each plan gets back only its own data, the effort is led by a group outside the plans, and there is no direct cooperation about the structure of the actual payouts, the likelihood of antitrust or other law suits is minimal.

If direct cooperation is not possible among plans or employers in a given region, an individual plan or employer can still successfully employ P4P by limiting its focus to a small number of high-volume providers and relying on an outside source of information and benchmarking of providers. As noted, a number of P4P programs such as BTE use practice reported medical chart abstract data reviewed by an external entity (in the case of BTE, NCQA fulfills this role through use of its recognition programs) to determine eligibility for bonus payments. Even in these situations, it is important that plans or purchasers use a common set of measures and feedback and coordinate their efforts with other plans and purchasers, as well as with providers, to ensure success of the P4P program effort.

Summary. In summary, there are multiple pathways and measurement approaches open to developing P4P programs. If done well, P4P appears to be a tool that offers a constructive path to aligning the incentives and interests of employers, plans, and providers, and takes ongoing improvement in quality and efficiency from the realm of the theoretical to being an integral part of our healthcare system. ■

REFERENCES

1. Spikoff M. Will pay for performance programs introduce a new set of problems? *Manag Care*. 2004; 13:16-17, 21-22.
2. Endsley S, Kirkegaard M, Baker G, et al. Getting rewards for your results: pay-for-performance programs. *Fam Pract Manag*. 2004;11: 45-50.
3. HEDIS—Health Plan Employer Data and Information Set. Washington DC: National Committee for Quality Assurance; 2004.
4. CAHPS—Consumer Assessment of Health Plans Study. Washington, DC: National Committee for Quality Assurance; 2004.
5. Bridges to Excellence. NCQA Physi-

cian Practice Connection. http://www.bridgestoexcellence.org/bte/physician/pol_recog_practice.htm. Accessed 9/20/04.

6. PriceWaterhouseCoopers. Pay for Performance's Small Steps of Progress. 2004. <http://www.pwc.com/healthcare>. Accessed 9/30/04.
7. National Committee for Quality Assurance. ADA-NCQA Diabetes Recognition Program. <http://www.ncqa.org/DPRP/dprpmain.htm>. Accessed 9/20/04.
8. National Committee for Quality Assurance. AHA-ASA-NCQA Heart Stroke Recognition Program. <http://www.ncqa.org/hsrp/hsbroche.pdf>. Accessed 9/20/04.
9. Harris Interactive. 2001 International Health Policy Survey. <http://www.cmwf.org>. Accessed 9/24/04.
10. Poe SL. Companies unite in care focused purchasing initiative to push US new health care market. http://www.mercerhr.com/press_release/details.jhtml?idContent=1132360. Accessed 9/20/04.
11. Integrated Healthcare Association. History of IHA's Pay for Performance Initiative. <http://www.iha.org/payfprfd.htm>. Accessed 9/20/04.
12. Massachusetts Health Quality Partnership. <http://www.mhqp.org/>. Accessed 9/20/04.