ABSTRACT

BACKGROUND: Pay for performance (P4P) initiatives are designed to foster and reward improvement in health care delivery. These programs promote “value-based health care” by rewarding quality care that is characterized by a reduced amount of disproportionate spending.

OBJECTIVE: To review the intent and design of P4P initiatives as well as the design and results of P4P programs in current practice.

SUMMARY: Three key principles are fundamental to building a value-based health care system: measurement, transparency, and accountability. There are several levers currently driving P4P, each influencing the movement in its own way. Among these are employers, federal agencies such as the Centers for Medicare & Medicaid Services and the Department of Health and Human Services, health plans, providers, accreditors, and Congress. One key player in the P4P movement, the National Committee for Quality Assurance (NCQA), is a private, independent nonprofit health care quality oversight organization that measures and reports on health care quality and unites diverse groups around a common goal: improving health care quality. NCQA, has demonstrated several successful provider-level measurement initiatives connected to P4P programs, notable among them Bridges to Excellence programs in several markets, physician recognition programs, the Integrated Healthcare Association’s P4P initiative in California, the National Forum on Performance Benchmarking of Physician Offices and Organizations, and health plan accreditation.

CONCLUSIONS: The initial data from developmental P4P programs across the nation have indicated that both financial and nonfinancial incentives motivate significant change in health care delivery, but the return on investment of these initiatives is not yet known.

KEYWORDS: P4P, NCQA, HEDIS, BTE, IHA, Measures, Transparency, Accountability, Value-based health care, Incentives

J Manag Care Pharm. 2007;13(2)(suppl S-b)S3-S6

The pay for performance (P4P) movement in managed care is based on the premise of fostering and rewarding improvement in health care. Most P4P programs are constructed to promote “value-based health care” by rewarding the types of care that tend to reduce the need for the disproportionate spending on care of patients with severe disease or high risk for disease. “Value-based” health care ultimately promotes a more effective distribution of funds and efforts across the continuum of health through measurement, transparency, and accountability. P4P programs should be rooted in guidelines established from clinical evidence and developed in collaboration with purchasers, policymakers, plans, and the providers themselves.

NCQA and Pay for Performance

The National Committee for Quality Assurance (NCQA) is a private, independent nonprofit health care quality oversight organization. The organization measures and reports on health care quality and unites diverse groups around a common goal: improving health care quality. NCQA’s mission is to improve the quality of health care, while the organization’s vision is to transform health care through measurement, transparency, and accountability.

NCQA has demonstrated several successful provider-level measurement initiatives connected to P4P programs; notable among them are Bridges to Excellence programs in numerous markets, physician recognition programs, the Integrated Healthcare Association (IHA) P4P initiative in California, the National Forum on Performance Benchmarking of Physician Offices and Organizations, and health plan accreditation.

NCQA physician recognition programs serve as an excellent case study for the P4P movement. They demonstrate the value of measuring structure, process, and outcomes, and rewarding high performance appropriately.

Initial results from such P4P programs strongly suggest that measurement coupled with rewards leads to improvement in health care quality. Furthermore, physicians appear to be the most crucial group in the engagement and implementation of P4P initiatives, with a focus on actionable performance feedback. While the success demonstrated by these programs is encouraging for the managed care community, only time will tell if the overall movement will, in fact, be fiscally sound, as payers’ return on investment (ROI) of P4P programs becomes clearer.

Improving Health Care Quality With P4P

In 2001, the Institute of Medicine (IOM) recognized a void in the current state of health care quality and released a landmark report, Crossing the Quality Chasm: A New Health System for the 21st Century. In this report, the IOM outlined several suggested initiatives for improving health care quality: creating an infrastructure to support evidence-based practice, facilitating the use of information technology (IT), preparing the workforce to better serve patients in a world of expanding knowledge and rapid
Measurement in P4P

Three key principles are fundamental to building a value-based health care system: measurement, transparency, and accountability. These principles are presented in deliberate order. Measurement is the primary concept in improving the quality of health care; without measurement, there is no baseline level of quality against which to gauge improvement. Transparency serves 2 purposes: first, it ensures that quality data is translated into measures and reports that consumers and purchasers can understand and use to make informed decisions. Second, valid measurement is crucial to assigning accountability: those who deliver health care cannot be held accountable for improvement if the quality of care itself is not accurately measured.

Metrics employed in P4P initiatives should be evidence-based and reflect a consensus of key stakeholders. P4P metrics come in 2 forms: measures and guidelines. Measures indicate desired clinical outcomes. For instance, a measure might assess whether a physician’s patients with cardia conditions control their low-density lipoprotein cholesterol (LDL-C) to 100 mg/dL or less. On the other hand, guidelines monitor evidence-based “best-practices” that lead to that desired result; a related guideline would ask whether the same doctor ordered annual cholesterol tests for his cardia patients. Measures and guidelines alike are useful as measurement tools in P4P programs.

A useful P4P measure is relevant, sound, and feasible: measures that lack these qualities are a waste of time and resources. Relevant measures will be designed to have a significant impact on health outcomes; advance cost-effective, evidence-based methods of care; have strategic significance to practitioners and health plans; and will assess aspects of care that can be improved. Measures must also be sound; that is, specific enough to reliably provide accurate assessments of care. Only trustworthy, evidence-based measures will be embraced by health plans, providers, consumers, and purchasers. Feasible P4P measures are practical for use in the “real world” of health care: they are characterized by precise specifications, a reasonable cost of measurement, and assured confidentiality. Feasible P4P measures also promote accountability by being open to a third-party audit, such as by a health plan or other P4P program sponsor.

The sorts of P4P programs that are possible are driven by 2 factors: the level where measurement takes place—the plan, hospital, or physician office—and the availability of relevant data. Take, for instance, the measurement of pharmacy claims data. Over the short term, limited initiatives at the plan or physician level may arise from pharmacy claims data taken in isolation. However, a more comprehensive pharmacy P4P program would require that those same pharmacy claims be linked to diagnoses or other data that may not be widely collected at the present time.

On the plan level, measurement has led to improved health care, specifically for certain Health Plan Employer Data and Information Set (HEDIS) measures developed by NCQA. In the 5 years from 1999 to 2004, NCQA documented an average increase of more than 52% for HEDIS effectiveness-of-care measures for chicken pox vaccination, hypertension, LDL-C control, and asthma (Figure 2). A useful P4P measure is relevant, sound, and feasible: measures that lack these qualities are a waste of time and resources. Relevant measures will be designed to have a significant impact on health outcomes; advance cost-effective, evidence-based methods of care; have strategic significance to practitioners and health plans; and will assess aspects of care that can be improved. Measures and guidelines alike are useful as measurement tools in P4P programs.

Similar improvement has also been observed on the physician level, where measurement has also led to advances in clinical quality. For example, performance among applicants to the Diabetes Physician Recognition Program (DPRP) showed substantial improvement in key measures, such as glycosylated hemoglobin (A1C) control, blood pressure control, and lipid control. 3

Many Voices in the P4P Debate

Entities ranging from individual physicians to federal agencies have waded into the P4P fray, each influencing the movement in its own way. Employers, health plans, accreditors, and Congress
have also weighed in. The Centers for Medicare & Medicaid Services (CMS) and the Department of Health and Human Services (DHHS) influence providers through demonstration projects and payment updates. Employers in several markets have banded together to sponsor programs that provide financial incentives to high-performing physicians and practices. Likewise, health plans support provider-performance measurement and recognition programs; while providers and consumers alike benefit from transparent ratings. Accreditors further influence the P4P movement by creating new evaluation tools for measurement and reporting.

Federal agencies impact P4P on the physician level primarily through CMS’s Physician Voluntary Reporting Program (PVRP). This program uses several different methods for collecting data from physicians’ offices, including claims for Current Procedural Terminology (CPT) Category II codes and G-codes. NCQA is assisting in the development of a validation methodology for the program through confidential feedback reports. Other Medicare demonstration projects such as the Physician Group Practice Demonstration, which involves the coordination of Part A and Part B services, also exist. DHHS also provides grants to build electronic health systems, including $18.6 million for 12 regions to link doctor offices, clinics, and hospital networks using open data standards by the end of 2006.4

The impact of health plans and employers on P4P can be seen in collaborative efforts such as the IHA and Bridges to Excellence programs in which NCQA plays a key role, as well as in the Massachusetts Health Quality Partners and Minnesota Community Partnership. Conversely, single-plan initiatives exist on this level. Notable among them are those developed by the Excellus/Rochester IPA and BlueCross/BlueShield of Michigan; the latter program demonstrated improved cardiac care and a 45% reduction in infection rates.5

New NCQA voluntary standards for health plans further influence these efforts through the promotion of standardized measurement and reporting of physician and hospital performance results. The new Physician and Hospital Quality (PHQ) standards have found robust support in the market; the program has earned 37 employer and consumer endorsements and 49 plans are early adopters of PHQ.6 Plans that voluntarily participate are required to use standardized measures, provide transparency about measurement, share measurement with those being measured, collaborate with other plans contracting with the same providers, and use results for reporting and for other quality improvement (QI) activities.

California P4P— IHA

One of these programs—IHA’s P4P initiative in California—has demonstrated both longevity and success in the P4P arena. IHA comprises 7 health plans with 6.2 million commercial managed care organization (MCO) enrollees and 225 capitated medical groups that provide for multiple health plans.7 There are 35,000 physicians in IHA, and Kaiser Permanente joined the association at the end of 2006.8 In this landmark collaboration, health plans and medical groups agree on measures, and plans combine their data. Clinical data is administrative only in IHA, and patient experience data surveyed at the group level. IT adoption is a performance standard employed by IHA that is evaluated by NCQA; NCQA also serves as the data aggregator. Based on the results of this monitoring, health plans make individual decisions on rewards with P4P recommendations.

The clinical measures employed by IHA have evolved over the past 3 years, from 6 measures in 2003 to 10 measures in 2005: childhood immunizations, cervical cancer screening, breast cancer screening, asthma management, A1C screening, A1C control, LDL-C screening among patients who had a cardiac event, LDL-C control <130 mg/dL for patients who had a cardiac event or were diagnosed with diabetes, chlamydia screening, and appropriate treatment for children with upper respiratory infection.9

Looking at the third-year results—that is, results reflective of care delivered in 2005—significant improvement was demonstrated in 3 categories: breast cancer screening (64% of physician groups improved in year 3), cervical cancer screening (61% improved), and diabetes: A1C screening and control (54% improved screening).

A particularly notable result arising from the third-year data was an average improvement of 40% among those physician groups who achieved a full IT score.10 This suggests that adoption of IT facilitates gains in quality. The relationship between IT adoption and clinical quality is most prominently displayed among groups scoring between 0% and 20% of the IT score. Groups who more fully integrated technology into their delivery systems tended to post higher clinical quality scores.10
NCQA Recognition Programs

NCQA sponsors 3 physician recognition programs, either alone or in conjunction with leading national health organizations: NCQA Practice Connections, including 1,600 physicians in 200 practices; NCQA/American Diabetes Association (ADA) Diabetes, including 2,600 physicians; and American Heart Association (AHA)/American Stroke Association (ASA)/NCQA Heart/Stroke, including 400 physicians. Measures included in these programs include structure, process, and outcomes of excellent care management. More than 3,800 physicians are recognized nationally through these programs, and they are rewarded by many health plans and Bridges to Excellence employers. To earn recognition, physicians must achieve certain standards of care across their entire panel of eligible patients. In the DPRP, for example, the total weight of all the scored measures is 100; physicians must achieve 75% to receive recognition. Between 2003 and 2005, there was a 33% increase in the number of DPRP-recognized physicians. In Bridges to Excellence Diabetes Care Link areas, the increase was particularly striking—the number of recognized physicians increased 450% over the same time frame.

Conclusions

The initial data from fledgling P4P programs across the board has indicated that while financial incentives do indeed motivate significant change, nonfinancial support also promotes quality improvement. The engagement of physicians is critical to the P4P movement, and public reporting heightens physician awareness. Experience has shown that actionable feedback on performance is a prerequisite to gains in clinical quality. Data integrity increases trust in the results.

NCQA’s data on the first wave of P4P programs has indicated that measurement provides physicians with a new perspective on their practice and that practices change their processes and delivery systems in order to meet program standards. National standards appear to be just as difficult to achieve for small and large practices, but having national measures helps reward programs get started. Physicians (i.e., generalists and some specialists) appreciate consistent requirements for measures. Clinical data continues to be difficult to obtain, as chart abstraction, especially in the absence of widespread use of electronic medical records, requires a significant commitment of time and resources.

P4P is not a silver bullet to cure all ills; it is a useful tool to align payment incentives. How a P4P initiative is conducted is just as important as whether it is conducted—it must be based on widely recognized, evidence-based measures and developed in collaboration with the providers it proposes to measure. Measurement and payment functions must be kept separate, and payers are right to be skeptical of additive payments.

The bottom line is that measurement plus rewards equals improvement in health care quality. While initial P4P programs have demonstrated success in improving health care quality, the financial implications of P4P are not yet known since the ROI data pertaining to such programs is incomplete at this time.

DISCLOSURES

This article is based on a presentation given by the author at a symposium, “Pay for Performance: Where’s the Return?” held October 4, 2006, at the Academy of Managed Care Pharmacy’s 2006 Educational Conference in Chicago, Illinois. The symposium was supported by an educational grant from Merck & Co., Inc. The author discloses that she has received an honorarium from Merck & Co., Inc. for participation in the symposium and this supplement. She discloses no potential bias or conflict of interest relating to this article.

REFERENCES