

A Pay-for-Performance Program for Behavioral Health Care Practitioners

Anthony L. Pelonero, M.D.

Richard L. Johnson, M.S.H.A., R.N.

This column describes a pay-for-performance program for behavioral health care practitioners. Implemented in 1996 by a large national health insurer, the program's goals are to improve the quality of care, recognize the practitioners who provide higher-quality care, demonstrate the value of behavioral health services to purchasers, and help providers align their practices with national standards. A future goal is to provide patients with data on provider quality to improve their treatment decisions. Important considerations in measure development include application of the measure to all disciplines, feasible data collection processes for providers, creation of clinically meaningful and fair measures, and selection of measures with large baseline variability. *Psychiatric Services* 58:442–444, 2007

Public and private health care payers are using pay-for-performance programs to provide incentives for practitioners to work toward improving the quality of care. The Centers for Medicare and Medicaid Services has several pay-for-performance initiatives (1). This linkage of reimbursement and quality has met with reaction from

professional organizations. For example, the American Medical Association has recommended principles for such programs (2). A handful of medical specialties have developed specific quality and outcome measures that are entered into national databases (3–5). Compared with other fields, psychiatry and other behavioral health care disciplines are early in the process of developing quality measures.

The Institute of Medicine (IOM) has published recommendations to improve the quality of mental health and substance abuse care (6). Many of the recommendations are specifically directed at health plans and direct payers for treatment services; specific recommendations for direct providers of mental health and substance abuse services are also included. Recommendations such as “use quality comparisons when making purchasing decisions” and “provide consumers with comparative information on the quality of care provided by practitioners” may have substantial impacts on practitioners if implemented by payers on a large scale.

Anthem Blue Cross and Blue Shield of Virginia implemented a pay-for-performance program for primary care practitioners in 1999. The program rewards primary care physicians a bonus payment for meeting various measures that reflect high-quality care. Several physician specialties and hospitals have been added to our pay-for-performance initiatives in the past several years. The Anthem behavioral health care pay-for-performance program is one of the few in the country specific to behavioral health care practitioners. This column provides an overview of the program and the challenges of implementing it.

Anthem's pay-for-performance program

The Anthem program began in 1996 as a “quality incentive” program for our commercial health maintenance organization (HMO) line of business. It was initially funded from savings realized when costs of care were lower than the predicted budget across the covered population. When the plan had sufficient experience to accurately predict annual cost of care, savings were accrued to support the program. In 2005 the program expanded from our HMO business to include other products, such as preferred provider organizations (PPOs).

A number of goals form the rationale for the program. The first goal is to measure quality for the purpose of trying to improve behavioral health care processes and outcomes. The second goal is to allow the differential reimbursement of providers; providers that are providing higher quality of care are better compensated. The third goal is to demonstrate the value of mental health and substance abuse services so that they are not viewed as a commodity by purchasers. Fourth, in the past few years health plans have had to show actual improvement in measures on the Health Plan Employer Data and Information Set (HEDIS), and our pay-for-performance measures have evolved to align with the quality performance objectives of the plan.

Participation in the pay-for-performance program is voluntary and open to all independently licensed behavioral health care disciplines in all of our lines of business. Very few eligible provider groups have expressly declined to participate (3.4%). To be eligible, a provider group must meet a minimum threshold for patient vol-

The authors are affiliated with Anthem Blue Cross and Blue Shield of Virginia, Mail Drop VA4003-M000, 2221 Edward Holland Dr., Richmond, VA 23230 (e-mail: anthony.pelonero@anthem.com). Dr. Pelonero is also a clinical associate professor at Virginia Commonwealth University in Richmond. Steven S. Sharfstein, M.D., Haiden A. Huskamp, Ph.D., and Alison Evans Cuellar, Ph.D., are editors of this column.

ume so that participating practices will have large enough samples to allow measurement of most of the pay-for-performance measures.

Performance data are aggregated by tax identification number to aid in administration of the program and to obtain the necessary sample sizes. Data collection methods are specific to each measure and include medical record audit sampling, review of administrative claims data, and member surveys.

Measures are first developed and reviewed by our management team and then vetted with a behavioral health care provider advisory committee. After specifications are agreed upon, test samples are reviewed to establish baseline performance and incentive targets. The plan then conducts several face-to-face meetings with practice representatives to review the recommended measures. Participating provider groups are exposed to the measures for at least one year before performance on the given measures determines any portion of eligibility for financial award. During this initial period, they are given data on their baseline performance as it compares with the rest of the network.

We have encountered many challenges in the development of pay-for-performance measures specific to behavioral health. We have identified six key points in the design of such measures. First, measures must apply to all or most behavioral health practitioners regardless of discipline or licensure. We did not limit our program to physician-only measures because in our plan psychiatrists constitute 20% or less of behavioral health practitioners and 29% of outpatient behavioral health specialty claims.

Second, it is important that data be reasonably obtainable. Burdensome data collection, such as record review, is often unacceptable and cost-prohibitive. Querying an administrative data base is preferable. However, certain limitations must be recognized and accepted when administrative data are used.

Third, aggregating and reviewing data by diagnosis often leads to inadequate sample sizes; grouping patients by diagnostic categories, or families, is more likely to yield meaningful samples. For example, the

largest portion of our outpatient claims (approximately 40%) fall into the depression diagnostic group, which led us to add a depression-specific process-of-care measure.

Fourth, measures should be clinically meaningful; not only should they reflect good-quality care but they should also take into account the realities of actual practice.

Fifth, the vetting process for the measure will reveal whether providers view it as fair. Providers will indicate whether they—or their peers—will be at a disadvantage on a given measure on the basis of their type of practice. For example, measures that are limited to adults will not capture providers who see primarily children.

Sixth, baseline rates for measures under consideration should optimally show wide variability (a large standard deviation), because one of the indicators of improvement is decreasing variation. Obviously, as the standard deviation grows smaller, decreasing variation is more difficult to accomplish; thus selecting measures with a small standard deviation at the baseline evaluation makes future gains less likely.

Over the past ten years selection of pay-for-performance measures has proceeded slowly but steadily—from simple measures of process (for example, whether the provider asks about a substance abuse problem) to outcomes of the provider's intervention (for example, whether the patient attended further sessions or a substance abuse treatment program). Our measures, whether process or outcome oriented, are typically supported by guideline-recommended care as well as by measures developed by national organizations that promote quality improvement. HEDIS represents the national standard of quality of care for health insurance plans, and our goal of improving quality led us to select pay-for-performance measures that align more directly with HEDIS. Additionally, we have found that basing measures on HEDIS promotes providers' acceptance, because the measures are based on the opinions of national clinical experts and thought leaders.

Measures typically target the high-volume and high-risk behavioral health populations (such as patients with major depression and substance

use disorders and children) and recognized care deficiencies identified through plan operations (such as identifying and addressing a substance abuse problem). [A table listing examples of measures we have used or are currently using is available in an online supplement to this column at ps.psychiatryonline.org.]

Each measure is weighted as a percentage of the group's total possible reward. If a group is unable to report data for some measures, the value of those measures is collapsed into the measures with adequate samples.

There is a constant balancing act between ensuring that the measures meet the considerations discussed above and that they are sufficiently complex and clinically meaningful to remain palatable to both providers and the plan and to address the key elements of care provided to the members.

Providers' reactions and program results

We have found that there will always be some amount of provider pushback in response to the program or some aspect of it, but pushback has never been a deal-breaker. Initially the program had no "down side"—that is, providers were paid by the standard fee schedule and their participation in the incentive program did not put compensation at risk. More recently, fee schedule increases have been rolled into the pools of incentive dollars. Demonstrating that the potential reward of participation is greater than a simple across-the-board fee schedule increase has been a helpful selling point with practitioners. We frankly explained that the money earmarked for a fee increase was focused on the providers that were most important to the plan—those in the network with high volume. Although volume itself is not an inherent value, it gains importance because of the number of members whose care can be affected.

As emphasized by Gosfield (7), our experience has been that engaging providers and developing a trusting relationship is a key ingredient in implementing and sustaining a pay-for-performance program. There is no shortcut to a collegial approach and practitioner involvement. Input to program design, provision of raw data, and

open, consistent, and candid communication cannot be overemphasized. In hindsight, some of our success can also be attributed to two factors. First, we started the program with a managed care-savvy HMO network and several years later expanded it to our larger PPO provider network. Second, the pay-for-performance program is conducted in the context of a managed behavioral health care program that has used profiling data to practitioners' advantage to eliminate plan management for most practitioners (for example, prior authorization or treatment plan submission is not required); the program is conducted with an underlying value of measuring quality to prevent mental health care from being viewed simply as a commodity.

Several stakeholders have informed us that the program is meaningful to them. Providers in the plan who see a significant volume of our members indicate that the financial reward has promoted their engagement with the plan administrators and increased their interest in the measures. The health plan's sales force indicates that measuring the quality of care provided by network providers promotes the value of the behavioral health care benefit. The value of the program as an educational tool for members is evident with the growth of consumer-directed health plan products. A pending challenge will be how to provide consumers with the identity of participating providers and the measurement results in a fair and meaningful way.

We have results from the program on six measures, which have been tracked from three to eight years. [A table showing changes over time is available in an online supplement to this column at ps.psychiatryonline.org.] Comparison of the current rate to the baseline rate indicates that improvement has been achieved on all measures except one. Although we have not yet attempted to validate the differences over time statistically, a preliminary conclusion that our pay-for-performance program has resulted in higher quality of care does not seem unreasonable. The single measure that has not improved is our patient satisfaction survey, which rates the likelihood that the patient would recommend the provider to a friend or col-

league. In 2007 we will be making comments and individual provider scores available to the groups. Not all improvements in rate have been large. However, we expect that our gains will be slow and sometimes small, given the large denominators for the network-level rates and the temporal measurement windows (12 to 24 months for some rates).

Future directions

Pay-for-performance programs are growing in the United States, primarily driven by health care financing organizations. The primary goal of such programs is to link financial compensation to the delivery of safe, effective, evidence-based health care.

We believe our measures are consistent with the IOM's recommendations, although they were developed before the publication of the IOM report (6). In a behavioral health population, risk adjustment has been particularly difficult, particularly because of the variability observed in diagnosis and in coding of claims.

A transparent and collaborative process of measure development and data sharing has been critical to ongoing acceptance of our program. Continued engagement of practitioners will be essential as measures move from process to outcomes, especially as more practitioners participate and have their compensation linked to results.

As we have noted, there are several reasons to implement a pay-for-performance program, primary among them the twin goals of improving the quality of care and differentially rewarding high-quality care. A third reason—one that points toward what we believe to be the future of such programs—is to develop a set of tools that can be used by our members to optimize use of their mental health and substance abuse benefit. If we can provide members a set of tools by which to differentiate between higher-quality and lower-quality care, we then enable them to self-select for higher-quality care, achieving our primary goal of ensuring the highest possible quality of care for the greatest possible percentage of members. Members often wish to consider cost in making health care choices. The provision of cost data as a supplement

to quality data is another important area for consideration.

We also hope to have an impact on insurance purchasing behavior of individuals and employer groups. A network built on the concept of large, key groups that participate in quality measurement and pay-for-performance programs is a marketable advantage. By providing pay-for-performance metrics, plans help purchasers move from decisions based primarily on cost to those based on both cost and quality data. Once quality becomes a purchasing consideration, behavioral health care cannot be viewed as a commodity.

Using pay-for-performance metrics to enhance selection may in the end prove a more powerful inducement to practitioners to work to improve quality than purely financial incentives. Much work remains to be done to establish the optimal formats and venues for the use of such metrics as a tool for consumer-directed health care.

Acknowledgments and disclosures

The authors thank the practitioners participating on Anthem's Behavioral Health Advisory and Oversight Committees for their collaboration.

Other than their employment by Anthem Blue Cross and Blue Shield of Virginia, the authors report no competing interests.

References

1. Medicare "Pay for Performance (P4P)" Initiatives. Baltimore, Centers for Medicare and Medicaid Services. Available at www.cms.hhs.gov/apps/media/press/release.asp?counter=1343
2. Principles for Pay-For-Performance Programs, 2005. Chicago, American Medical Association. Available at www.ama-assn.org/ama1/pub/upload/mm/368/principles4pay62705.pdf
3. STS National Database, 2006. Chicago, Society of Thoracic Surgeons. Available at www.sts.org/sections/stsnationaldatabase
4. National Surgical Quality Improvement Program, 2006. Chicago, American College of Surgeons. Available at <https://acsnisqip.org/main/abouthistory.asp>
5. National Cardiovascular Data Registry. Washington, DC, American College of Cardiology Foundation. Available at www.acc.ncdr.com/webncdr/common
6. Institute of Medicine Committee on Crossing the Quality Chasm: Improving the Quality of Health Care for Mental and Substance-Use Conditions. Quality Chasm Series. Washington, DC, National Academies Press, 2005
7. Gosfield AG: P4P: transitional at best. *Managed Care Magazine*, Jan 2005. Available at www.managedcaremag.com/archives/0501/0501.p4pgosfield.html