Medicare Physician Group Practice Demonstration Design: Quality and Efficiency Pay-for-Performance

John Kautter, Ph.D., Gregory C. Pope, M.S., Michael Trisolini, Ph.D., M.B.A., and Sherry Grund, R.N.

The Medicare Physician Group Practice (PGP) demonstration is Medicare's first physician pay-for-performance (P4P) initiative. The demonstration, which is legislatively mandated, establishes incentives for quality improvement (QI) and cost efficiency at the level of the PGP. Ten large physician groups are participating in the demonstration, which started on April 1, 2005, and will run for 3 years. In this article the authors provide an overview of the PGP demonstration's key design elements, including the selection process for PGP participants; beneficiary assignment; comparison population; measurement of demonstration savings; performance payments; and quality measurement and reporting. A summary of early case study findings is also provided.

INTRODUCTION AND OVERVIEW

The Medicare physician fee schedule was established as part of the 1989 Omnibus Budget Reconciliation Act. In addition to establishing a standardized payment schedule based on a resource-based relative value scale, it established a physician payment formula based on achievement of an expenditure target—the volume performance standard (VPS). However, the VPS approach had several methodological flaws that prompted its replacement with the sustainable growth rate (SGR) system in the 1997 Balanced Budget Act. Nonetheless, the SGR approach has run into difficulties as well, including volatile updates that in some years have been too high and in others too low (Medicare Payment Advisory Commission, 2005).

The SGR, and its predecessor the VPS, are budgetary tools, but they do not establish strong incentives for providers to slow volume and intensity growth. These approaches have several key weaknesses. They are national targets, which dilute incentives so as to be barely noticeable for any individual provider. Providers who restrain volume are treated the same as those who do not. Also, their national scope means they do not take into account regional or local variations in market conditions that may affect rates of expenditure growth. Finally, they apply only to physician expenditures, which are a minority of total Medicare expenditures, and do not give providers incentives to coordinate services along the continuum of care or to improve quality of care.

The Medicare PGP demonstration, which is Medicare's first physician P4P initiative, attempts to overcome these limitations by establishing incentives for QI and cost efficiency at the level of the PGP. A legislative mandate for the PGP demonstration was included in the Medicare, Medicaid, and State Child Health Insurance Program Benefits Improvement and Protection Act of 2000. It established several goals, including:

- Encouraging coordination of health care furnished under Medicare Parts A and B.

John Kautter, Gregory C. Pope, and Michael Trisolini, are with RTI International. Sherry Grund is with the Iowa Foundation for Medical Care. The research in this article was supported by the Centers for Medicare & Medicaid Services (CMS) under Contract Number 500-00-0024. The statements expressed in this article are those of the authors and do not necessarily reflect the views or policies of RTI International, Iowa Foundation for Medical Care, or CMS.

- Encouraging investment in administrative structures and processes for efficient service delivery.
- Rewarding physicians for improving health care processes and outcomes.

The PGP demonstration started on April 1, 2005, and will run for 3 years. In addition, calendar year 2004 is used as a baseline for cost and quality performance assessment.

The premise of the PGP demonstration is that PGPs can achieve higher quality and greater cost efficiency by managing and coordinating patient care. The physician groups participating in the PGP demonstration are engaged in a wide variety of care management interventions to improve the cost efficiency and quality of health care for Medicare fee-for-service (FFS) patients (Kautter et al., 2006). These interventions include: chronic disease management programs, high-risk/high-cost care management, transitional care management, end-of-life/palliative care programs, practice standardization, and QI programs. In addition, information technology, such as electronic medical records, patient disease registries, and patient monitoring systems, are being used by PGP participants to improve practice efficiency and quality of care delivered to patients, and to better understand the utilization of services by the Medicare FFS population. The PGP demonstration will test whether care management initiatives generate cost savings by reducing avoidable hospital admissions, readmissions, and emergency department visits, while at the same time improving the quality of care for Medicare beneficiaries.

The PGP demonstration employs a shared savings provider payment model in which savings in Medicare expenditures are shared between participating physician groups and the Medicare Program. In effect, this model is a hybrid between the FFS and capitation payment methods. Providers continue to be paid under FFS rules, and beneficiaries are not enrolled (i.e., they retain complete freedom of provider choice). However, participating physician groups are able to retain—through annual performance payments in addition to their FFS revenues—part of any savings in Medicare expenditures that they generate for their patients. This shared savings payment model gives participating providers a financial incentive to control the volume and intensity of medical services, such as exists under capitated payment. Moreover, a higher portion of savings is retained by physician groups the better their measured quality of care. In this way, incentives for both cost efficiency and QI are introduced into FFS payment. Because participating providers retain only part of the savings generated by reducing expenditures, incentives for underservice and risk selection are lower than under full capitated payment. Another difference from capitation is that the Medicare Program shares in any savings, benefiting from cost efficiency improvements and lowering government expenditures.

In this article we describe the design of the PGP demonstration (Pope et al., 2002; Kautter et al., 2004; Trisolini et al., 2005), which builds on the group-specific volume performance standards design (Tompkins et al., 1996; Wallack and Tompkins, 2003).[1] We first explain the PGP demonstration's key design elements, including the selection process for PGP participants; beneficiary assignment; comparison population; measurement of demonstration savings; performance payments; and quality measurement and reporting. We then provide a summary of early case study findings for the PGP demonstration. Finally, we conclude with a review and discussion of several key issues.

[1] An evaluation of the PGP demonstration is currently underway. At the time this article was written, no evaluation results were available, except for early case study findings, which are summarized in this article.

PARTICIPANT SELECTION PROCESS

As shown in Table 1, 10 physician groups, located across the Nation, are participating in the PGP demonstration. They were selected through a competitive process by CMS, based on organizational structure, operational feasibility, geographic location, and implementation strategy. Large PGPs were selected to ensure that participants would have the administrative and clinical capabilities necessary to respond to the PGP demonstration's incentives. Further, large PGPs treat sufficient numbers of Medicare beneficiaries so that the calculation of performance payments is statistically reliable. An additional requirement for participating PGPs to be multispecialty groups is consistent with the expectation that PGP demonstration participants must possess the capacity to respond to incentives by coordinating care across multiple provider types and sites of care. The participating PGPs all have at least 200 physicians, and together represent more than 5,000 physicians. They include freestanding group practices, components of integrated delivery systems, faculty group practices, and physician network organizations.

BENEFICIARY ASSIGNMENT

A key aspect of the PGP demonstration design is patient attribution, or beneficiary assignment. The intent of the PGP demonstration is to create an incentive for each participating PGP to coordinate and manage the health care of the beneficiaries assigned to it. A PGP's ability to coordinate and manage the health care of a beneficiary depends on: (1) the type of services the PGP provides to the beneficiary, and (2) the overall control the PGP has over the beneficiary's utilization of services. Because the PGP demonstration is a Medicare FFS innovation, there is no enrollment process whereby beneficiaries accept or reject involvement. Therefore, we developed a methodology to assign beneficiaries to participating PGPs based on utilization of Medicare-covered services. Extensive simulations of alternative assignment methods were conducted using multiple years

Table 1

Health Care Organizations Participating in the Physician Group Practice (PGP) Demonstration

Participant	Organizational Structure	Part of Integrated Delivery System	Includes Academic Medical Center	Service Area
Dartmouth-Hitchcock Clinic	Faculty/Community Group Practice	Yes	Yes	New Hampshire/Eastern Vermont
Billings Clinic	Group Practice	Yes	No	South-Central Montana/Northwestern Wyoming
Geisinger Clinic	Group Practice	Yes	No	Central-Northeast Pennsylvania
Middlesex Health System	Network Model	Yes	No	South-Central Connecticut
Marshfield Clinic	Group Practice	No	No	North-Central Wisconsin
Forsyth Medical Group	Group Practice	Yes	No	Northwest North Carolina
Park Nicollet Clinic	Group Practice	Yes	No	South-Central Minnesota
St. John's Clinic	Group Practice	Yes	No	South-Central Missouri/Northwest Arkansas
The Everett Clinic	Group Practice	No	No	West-Central Washington
University of Michigan Faculty Group Practice	Faculty Practice	Yes	Yes	Southeastern Michigan

SOURCE: John Kautter, Ph.D., Gregory C. Pope, M.S., Michael Trisolini, Ph.D., M.B.A., RTI International, and Sherry Grund, R.N., Iowa Foundation for Medical Care.

of historical data on several large PGPs. Characteristics of assignment methodologies that we examined included the type of services provided by the PGPs to beneficiaries (Part B physician services, evaluation and management (E&M) services, outpatient E&M services); the share of a beneficiary's utilization provided by a PGP (e.g., all, majority, plurality, any); and, the type of physician (primary care, specialist) treating the beneficiary.

We evaluated the alternative assignment methodologies on two criteria: provider responsibility and sample size. First, providers must believe that the numbers and types of services they provide mean that they have primary responsibility for the health care of beneficiaries assigned to them. Otherwise, PGPs may have difficulty responding effectively to the demonstration incentives, so the assignment methodology will lack face validity. Second, sample size is critically important for the statistical reliability of performance measurement. If the number of beneficiaries assigned to a participating PGP is too low, then cost and quality performance measurement may be unstable.

We concluded that a beneficiary should be assigned to a participating PGP if the PGP provided the largest share, i.e., the plurality, of outpatient E&M services to the beneficiary. A beneficiary who receives at least one office or other outpatient E&M service from a participating PGP during a given year is eligible for assignment to the PGP in that year. If the beneficiary received more of those services (as measured by Medicare allowed charges) from the PGP than from any other physician practice (group or solo), then the beneficiary is assigned to the PGP. Certain E&M services, such as emergency department visits, do not reflect the PGP's ability to manage and coordinate the health care of beneficiaries, and are not used in the beneficiary assignment methodology. Beneficiary assignment is redetermined after each year based on that year's utilization patterns. A result of this assignment methodology is that no beneficiary can be assigned to more than one PGP, preventing CMS from paying performance payments more than once when multiple PGPs serve overlapping patient populations.

For the assignment methodology previously outlined, our simulation analysis showed that (1) approximately 50 percent of beneficiaries that were provided at least one Part B physician service by the PGP during a year were assigned to the PGP, with groups with greater primary care orientation having more patients assigned; (2) approximately 50 to 70 percent of beneficiaries that received at least one outpatient E&M service from the PGP during a year were assigned to the PGP; (3) PGPs provided around 80 to 90 percent of the outpatient E&M services for their assigned beneficiaries; and, (4) PGPs generally retained approximately two-thirds of their assigned beneficiaries from one year to the next. While alternative assignment methodologies performed better on our sample size criterion (e.g., assignment based on Part B physician services) or on our provider responsibility criterion (e.g., assignment based on the majority of utilization), none of the alternative assignment methodologies performed better on both criteria. In addition, a prior study of several large PGPs concluded, based on physician interview results, that PGPs believed they had primary responsibility for the health care of patients to whom they had provided the plurality of outpatient E&M services (McCall et al., 1998).

The assignment methodology incorporates outpatient E&M services provided by specialists as well as by primary care physicians. One reason for this is that specialists (e.g., cardiologists) often are the principal primary care provider for elderly

and chronically ill patients, and it is reasonable to expect them to take responsibility for these patients. In addition, the assignment methodology provides an opportunity for specialists to take responsibility for assuring that their patients' primary care needs are being met even if the specialist is primarily treating a specific problem on a referral basis.[2]

For the PGP demonstration's base year, the number of Medicare FFS patients assigned to the 10 physician groups ranged from 8,383 to 44,609, and totaled 223,203. Overall for the 10 physician groups, the percentage of assigned patients that were female was 57.5 percent, dually eligible for Medicare/Medicaid was 13.3 percent, and age 85 or over was 10.3 percent. These distributions were broadly similar to the Medicare FFS population (Centers for Medicare & Medicaid Services, 2006).

COMPARISON POPULATION

The purpose of the comparison population in the PGP demonstration is to provide a benchmark for the cost control performance of the participating PGP. Specifically, to predict what the per capita expenditure growth of PGP assigned beneficiaries would have been in the absence of the PGP demonstration. Growth in per capita expenditures is influenced by local factors including changes in wages and other input costs, diffusion of new medical techniques and technologies, practice style variations, competition, population density, and characteristics of local medical care providers. For this reason, local Medicare beneficiaries not assigned to the participating PGP are a natural comparison population. With this comparison population, the participating PGP will earn a performance payment if it performs better than its local competitors, who face similar market conditions.

A convenient way to define a PGP's service area is by the residence location of beneficiaries assigned to it. This patient origin approach to service area definition has been widely studied and recommended for hospitals (Baker, 2001), and appears to apply equally well to identifying PGP service areas. Defining service areas by patient origin is an empirically based method that reflects the actual catchment area of each participating PGP. It does this more accurately than prespecified administrative units such as the county, metropolitan statistical area, or State, or than prespecified geographic radiuses, such as 15 miles from the practice location. The service areas of multilocation practices are accurately identified with the patient origin approach, and are automatically adjusted for changes in practice locations when the service area is annually updated.

As part of the PGP demonstration design process, we calculated several alternative patient-origin-based service areas for individual PGPs. The alternatives differed on the criterion for including counties in a PGP's service area, specifically what percentage of a PGP's assigned beneficiaries must reside in a county for it to be included in the service area. The alternatives were evaluated on the following criteria:

- Includes the areas where most of the PGP assigned beneficiaries reside.
- Sufficient comparison population sample size to provide a stable target growth rate computation.
- Service area is geographically compact and contiguous and has face validity.
- Minimizes data collection and computation burdens (by including fewer counties).
- Satisfies first four criteria for a wide range of sizes, types, and locations of

[2] The role of physician specialty in the patient attribution algorithm may be a fruitful subject for future analysis and possible refinement. One issue that will need to be investigated is the reporting of physician specialty information in Medicare billing data.

PGPs (e.g., urban, rural, large, small, PGPs drawing from a wide range of counties and those drawing from only a few counties).

We concluded that counties where at least 1 percent of a PGP's assigned beneficiaries reside should comprise its service area. These counties typically include 80 to 90 percent or more of a PGP's assigned beneficiaries and satisfy our other criteria. Each participating PGP's service area is defined for the base year and redefined for each performance year, and may differ between years to reflect changes in the location of the PGP's assigned beneficiaries.

The comparison population for the PGP is drawn from the PGP's service area. To ensure that the comparison population is similar to beneficiaries assigned to the participating PGP, the comparison beneficiaries must meet similar criteria, for example, a comparison beneficiary must have received at least one office or other outpatient E&M service. Beneficiaries assigned to a participating PGP in the current or any prior performance year, or beneficiaries that received any office or other outpatient E&M services at the PGP in the current performance year, are not eligible for the comparison population. These beneficiaries may be affected by the incentives provided by the PGP demonstration.

MEASUREMENT OF SAVINGS

Demonstration savings, termed Medicare savings, measures the cost savings impact of the PGP demonstration, and defines the pool of savings that the participating physician groups and the Medicare Program share. To calculate Medicare savings in a performance year, first the participating PGP's annual per capita expenditure target is calculated (all expenditures are on a per capita basis):

Target Expenditures = PGP Base Year Expenditures × (1 + Comparison Group Growth Rate).

Target expenditures in the demonstration are PGP-specific; they are based on each PGP's base year expenditure level. PGP base year per capita expenditures are calculated for beneficiaries assigned to the PGP in the base year. The comparison group growth rate is defined as the growth in per capita expenditures in the PGP's comparison population between the base and performance years. Both the PGP base year expenditures and the comparison group expenditure growth rate are adjusted for case-mix change between the base and performance years using a modification of the CMS hierarchical condition categories (HCC), or CMS-HCC, risk-adjustment model (Pope et al., 2004; Olmsted, Pope, and Kautter, 2006).

Medicare savings are computed as the difference between the per capita expenditure target and the PGP's per capita expenditures in the performance year (for beneficiaries assigned to the PGP in the performance year),[3] multiplied by the number of full-time equivalent (FTE) beneficiaries (person years) assigned to the PGP in the performance year:[4]

Medicare Savings = (Target Expenditures – PGP Performance Year Actual Expenditures) × FTE Assigned Beneficiaries.

This is a retrospective calculation, because neither actual nor target expenditures are known until after the end of the performance year.

[3] Performance year expenditures are annualized by dividing expenditures by the fraction of the year alive and enrolled in Medicare. Performance year per capita expenditures are then weighted by this fraction.

[4] To determine FTE beneficiaries, the fraction of the year each beneficiary was alive and enrolled in Medicare is calculated. FTE beneficiaries equal the sum of these fractions (i.e., equals number of person years).

PERFORMANCE PAYMENTS

Figure 1 shows the process of calculating performance payments in the PGP demonstration. For each PGP, the first step involves determining whether or not annual Medicare savings are more than 2 percent of target expenditures. The 2-percent threshold is used to account for normal variation in expenditures.[5] Given that observed expenditures represent a combination of PGP cost-saving performance and normal variation, the higher the threshold, the less likely it is that performance payments will be paid due to normal variation (undeserved performance payments), but also the less likely it is that performance payments will be paid due to PGP cost saving behavior (deserved performance payments). A simulation analysis showed that a 2-percent threshold represented a reasonable balance between paying deserved performance payments and not paying undeserved performance payments (Pope and Chromy, 1997).

If the PGP holds the expenditures for its assigned beneficiaries more than 2 percent below its target, it is eligible to earn a performance payment for that performance year (assuming there are no accrued losses from previous years). The net Medicare savings are calculated as the amount of annual Medicare savings more than the 2-percent threshold.

The sharing rate for net Medicare savings was set high enough to give PGPs sufficient incentive to participate in the demonstration, including rewards for improving and delivering high-quality care, while allowing for significant Medicare Program savings. Based on simulation analysis, we determined that an 80-percent sharing rate met these criteria. As shown in Figure 1, the net Medicare savings are divided, with 80 percent going to the PGP performance payment pool and 20 percent retained by Medicare as program savings. The PGP performance payment pool is then itself divided between a cost performance payment and a maximum quality performance payment. In performance year 1, the cost performance payment and maximum quality performance payment shares of the PGP performance payment pool are 70 and 30 percent, respectively. In performance year 2, the respective shares are 60 and 40 percent, and in performance year 3, the shares are 50 and 50 percent. This was done to gradually increase the importance of quality performance in the PGP demonstration. The actual quality performance payment is then determined, based on the percentage of the PGP demonstration's quality targets the PGP met in the performance year. If all of the quality targets are met, then the entire maximum quality performance payment is earned by the PGP. However, if some of the quality targets are not met, then a portion of the maximum quality performance payment is retained by the Medicare Program.

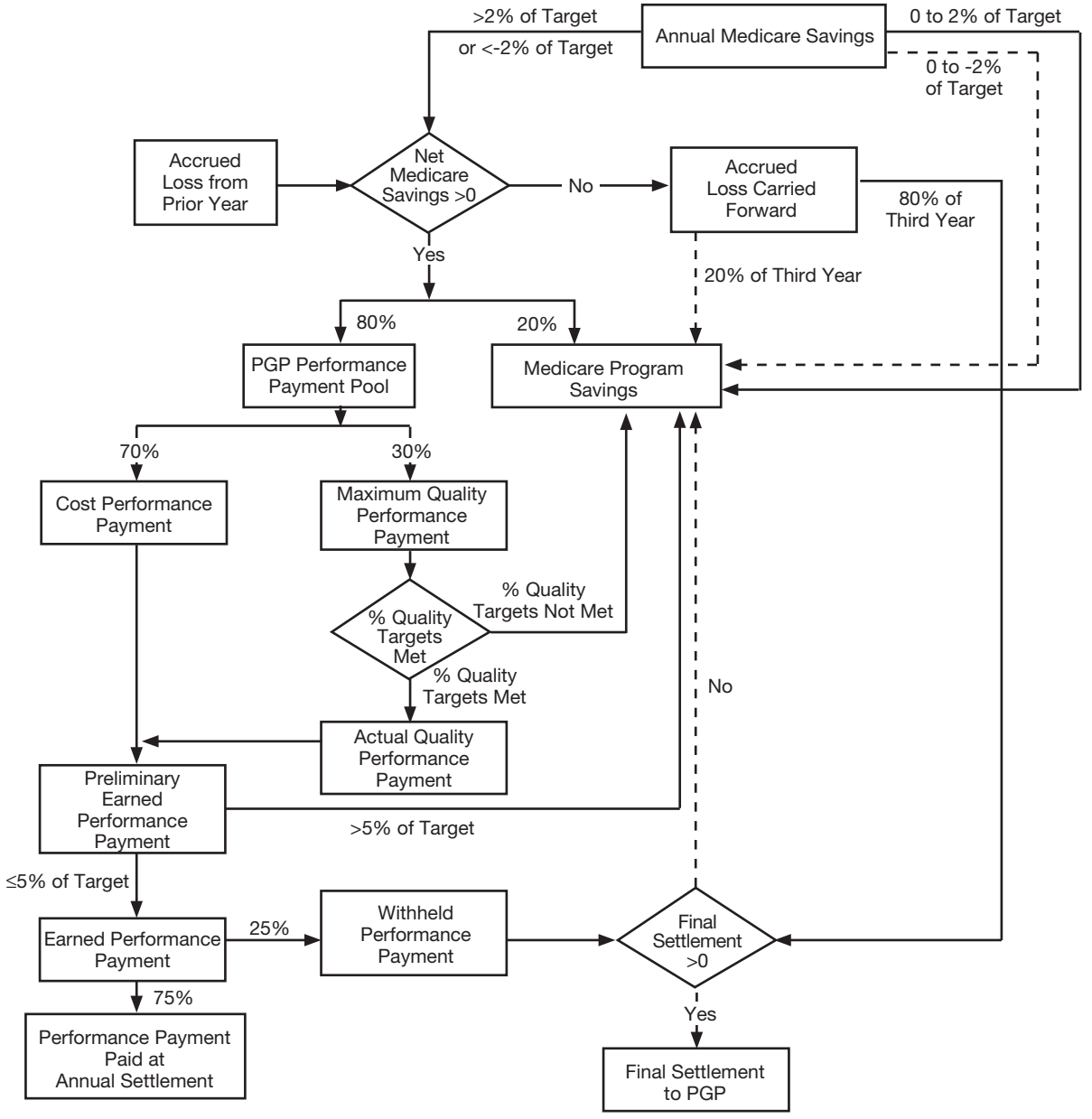
Once the actual quality performance payment has been determined, it is added to the cost performance payment to identify the preliminary earned performance payment, as shown in Figure 1. However, to avoid incentives for excessive cost cutting, the actual earned performance payment cannot be more than 5 percent of the PGP's target expenditures, which includes both Parts A and B expenditures; the final earned performance payment is capped at that 5 percent level if the preliminary earned performance payment is higher.

Finally, the performance payment paid to the PGP at the annual settlement will equal 75 percent of the earned performance payment amount. The other 25 percent of the earned performance payment will be withheld until the end of the demonstration

[5] Normal variation may arise from variations in the incidence of disease from year to year, or in claims-paying operations.

Figure 1

Process for Calculating Performance Payments in the Physician Group Practice (PGP) Demonstration



NOTES: Dotted lines represent negative contribution to Medicare Program savings. Annual Medicare savings between -2 and 2 percent of target expenditures are not included in performance payment computations because they may result from normal variation; they are included in Medicare Program savings. In performance year 1, the cost performance payment and maximum quality performance payment shares of the PGP performance payment pool are 70 and 30 percent, respectively. In performance year 2, the shares are 60 and 40 percent, respectively, and in performance year 3, the shares are 50 and 50 percent, respectively.

SOURCE: John Kautter, Ph.D., Gregory C. Pope, M.S., Michael Trisolini, Ph.D., M.B.A., RTI International, and Sherry Grund, R.N., Iowa Foundation for Medical Care.

to protect Medicare against losses the PGP may generate in subsequent years. At the final settlement, at the end of the demonstration, the cumulative amount of the withheld performance payments will be paid to PGP, after accounting for any accrued losses.

In a performance year, participating PGPs may perform worse than their comparison group and generate losses. That is, assigned beneficiary expenditures may exceed target expenditures, in which case Medicare savings are negative. Losses are defined as (negative) Medicare savings in excess of -2 percent of target expenditures. PGPs are protected against normal variation in expenditures between 0 and -2 percent. In addition, PGPs are not at risk to reimburse the Medicare Program for either annual losses or an accrued net loss at final settlement. However, annual losses are carried forward to the subsequent performance year and are used to offset (positive) Medicare savings generated in that year. No performance payment can be earned in a performance year unless Medicare savings are sufficient to offset accrued losses from prior performance years.

Annual Medicare savings between -2 and +2 percent of target expenditures generate neither losses to be carried forward nor performance payments to be paid (Figure 1). This portion of the annual Medicare savings (between – and + 2 percent) is assumed to be caused by normal variations in expenditure levels, not by the PGP's performance.

Cost savings are measured cumulatively from the original demonstration base year. Rebasing—meaning updating the base year for setting targets for the annual performance payment computation—does not occur. Not rebasing gives participating PGPs the maximum incentive to generate savings during the demonstration period. However, if the PGP demonstration model becomes part of the Medicare Program, periodic rebasing would be necessary to continue to provide incentives for improving the quality and efficiency of care and lock-in prior year savings so as not to indefinitely reward groups for prior performance.

QUALITY MEASUREMENT AND REPORTING

In this section we describe the quality measurement and reporting methods applied in the PGP demonstration. This includes quality (1) measurement, (2) targets, and (3) measurement processes and performance calculation. Both claims-based and medical record-based methods are used. Initially, the demonstration design included only claims-based quality measures. However, we worked with CMS to engage the physician groups participating in the demonstration to expand the quality measurement and reporting process to medical record-based measures which resulted in a consensus agreement for measuring and rewarding quality under the demonstration.

Quality

Measurement

Our overall approach had four main goals. First, we aimed to include a broad range of quality measures, so that participating PGPs would need to focus on a broad range of quality of care interventions and not just a select few. The danger with including only a small number of quality measures is that participating groups might focus excessively on them, to the detriment of other important quality of care objectives. This must be balanced against the added data collection and administrative burden imposed by including a wider range of measures. In addition, the number of measures should not be so broad that the incentive value of any individual measure becomes too weak to be an effective motivator.

The second goal was to use well established and validated quality measures, so that physicians practicing in the groups

Table 2

Quality Measures, by Module for the Physician Group Practice (PGP) Demonstration

Diabetes Mellitus	Heart Failure	Coronary Artery Disease	Hypertension and Preventive Care
DM-1 HbA1c Management	HF-1 Left Ventricular Function Assessment	CAD-1 Antiplatelet Therapy	HTN-1 Blood Pressure Screening
DM-2 HbA1c Control	HF-2 Left Ventricular Ejection Fraction Testing	CAD-2 Drug Therapy for Lowering LDL Cholesterol	HTN-2 Blood Pressure Control
DM-3 Blood Pressure Management	HF-3 Weight Measurement	CAD-3 Beta-Blocker Therapy – Prior Myocardial Infarction	HTN-3 Plan of Care
DM-4 Lipid Measurement	HF-4 Blood Pressure Screening	CAD-4 Blood Pressure	PC-5 Breast Cancer Screening
DM-5 LDL Cholesterol Level	HF-5 Patient Education	CAD-5 Lipid Profile	PC-6 Colorectal Cancer Screening
DM-6 Urine Protein Testing	HF-6 Beta-Blocker Therapy	CAD-6 LDL Cholesterol Level	—
DM-7 Eye Exam	HF-7 Ace Inhibitor Therapy	CAD-7 Ace Inhibitor Therapy	—
DM-8 Foot Exam	HF-8 Warfarin Therapy for Patients	—	—
DM-9 Influenza Vaccination	HF-9 Influenza Vaccination	—	—
DM-10 Pneumonia Vaccination	HF-10 Pneumonia Vaccination	—	—

SOURCE: John Kautter, Ph.D., Gregory C. Pope, M.S., Michael Trisolini, Ph.D., M.B.A., RTI International, and Sherry Grund, R.N., Iowa Foundation for Medical Care.

would accept them as appropriate. The third goal was to include quality measures primarily focused on ambulatory care, because that is the main setting for care provided by PGPs that focuses on coordination of care and prevention. Finally, the fourth goal was to focus on conditions highly prevalent in the Medicare population as well as conditions that account for a significant portion of Medicare spending.

A summary table that describes the PGP demonstration quality measures is presented as Table 2. The demonstration includes 32 quality measures covering four modules: (1) diabetes mellitus, (2) heart failure, (3) coronary artery disease, and (4) hypertension and preventive care. Each of the four modules includes quality measures applied only to beneficiaries with those specific diseases. However, the fourth module also includes measures applied to all beneficiaries that meet age and sex criteria. To demonstrate a high level of performance on quality of care for the demonstration, PGP participants will need to work actively to improve or maintain quality across a broad range of diseases and conditions.

The 32 quality measures are a subset of those developed by CMS' Quality Measurement and Health Assessment Group for the Doctors Office Quality (DOQ) Project (Centers for Medicare & Medicaid Services, 2005). As a result, they have been well established and validated through the extensive review process conducted as part of the DOQ project. The DOQ measures are also focused on care provided in ambulatory settings, which is emphasized in the PGP demonstration.

The 32 quality measures will be phased in so as to reduce the administrative burden faced by the PGP participants in collecting the medical records data needed for the 25 measures that require that type of data.[6] The schedule for phasing in the quality measures across the demonstration performance years is as follows:

[6] Further, the 32 quality measures are consistent with the clinical care guidelines for the chronic conditions that were available at the time of implementation of the PGP demonstration.

Performance Year 1—Diabetes measures, including flu and pneumonia vaccine measures for the diabetic population.
Performance Year 2—Year 1 measures plus the heart failure and coronary artery disease measures, including flu and pneumonia vaccine measures for the heart failure population.
Performance Year 3—Year 2 measures plus the hypertension measures and colorectal and breast cancer screening measures.

Targets

PGP participants are eligible to earn separate quality performance payments if they meet quality performance targets for each of the quality measures. For each measure, PGP participants must achieve at least one of three targets. The first two are threshold targets and the third is an improvement target:

- Achieve the higher of 75 percent compliance or the Medicare Health Plan Employer Data and Information Set (HEDIS®) mean for the measure (for those measures where HEDIS® indicators are also available).
- Achieve the 70th percentile Medicare HEDIS® level (for those measures where HEDIS® indicators are also available).
- Demonstrate a 10-percent or greater reduction in the gap between the administrative baseline and 100 percent compliance.

An example of how the improvement target is calculated is as follows. If a PGP achieves 40 percent compliance for a quality measure in the base year (2004), then the gap between that level and 100 percent is 60 percent. As a result, the PGP must reduce the gap by 10 percent of 60 percent, or 6 percentage points, so its QI target is 46 percent. If the PGP achieves 46 percent compliance with the quality measure in any of the three performance years of the demonstration, then it will be judged as having met the QI target for that measure for that year.

By including both threshold and improvement targets, participating groups are provided positive incentives for quality whether they start out at either high or low levels on measured quality for each indicator. If only threshold targets were included, then groups starting at low levels of quality might view the targets as unachievable. In contrast, if only improvement targets were included, then groups starting at high levels of quality might view further improvements as difficult to achieve.

Data Collection and Performance Calculation

Claims data analysis is used to calculate 7 of the 32 quality measures. They are given a weight of four in the overall performance calculation. The other 25 quality measures are calculated using only data from medical record abstraction or other internal PGP data systems. They are given a weight of one in the overall performance calculation.

The lower weight for medical record-based measures reflects the additional administrative burden to report those measures and the potential for larger variation because they are calculated from a random sample of 411 eligible beneficiaries. That sample size is adapted from the approach used for HEDIS® quality measures (National Committee for Quality Assurance, 2005). Because medical record review can be costly, a sampling approach is permitted for those measures. In contrast, the claims-based measures are calculated on all of the beneficiaries eligible for a given quality measure at each PGP, because the data required to compute those rates are available from existing claims data. They have larger sample sizes (for example, some PGPs have several thousand diabetics), the required data are

easily available, and their results are subject to less sampling variation; as a result, they receive higher weights in the overall performance calculation.

To calculate the overall quality performance for a given PGP, the number of quality measures for which the PGP has achieved either a threshold or improvement target is first calculated. Then the total weighted quality score is calculated by adding the value for each quality measure where a target was reached, either four or one, depending on the type of measurement. The weighted score for the PGP is then divided by the total possible score for the given performance year. The ratio is then applied to the quality portion of the performance payment pool to calculate the performance payment for quality (Figure 1).

EARLY CASE STUDY RESULTS

We conducted site visits to each of the 10 PGP participants during the first performance year of the PGP demonstration (Kautter et al., 2006). The purpose of these site visits was to understand the decisions of the PGPs to participate in the demonstration and their early implementation and operational experience with the demonstration. We interviewed demonstration sites about their reasons for joining the PGP demonstration, and their strategies for responding to its incentives. In general, physician groups report that their main reason for participating in the PGP demonstration is their interest in improving and managing patient care, their belief that this is the right thing to do for patients, and the alignment of the PGP demonstration with their mission and vision of the future of health care. Many of the participants have experience and infrastructure for care management from prior involvement with private insurers or Medicare managed care. They now wish to apply this orientation to the Medicare FFS population as part of their overall strategy of providing value to payers. They believe that participating in the PGP demonstration will position them to succeed in the future health care environment, which will reward the provision of high quality and efficient care.

PGP participants are implementing a variety of initiatives to improve the quality and efficiency of health care for Medicare FFS patients. These include chronic disease management, high-risk/high-cost care management, managing transitions between care settings (e.g., inpatient to outpatient), end-of-life/palliative care programs, and standardizing care around evidence-based protocols.

It is expected that care management programs will generate cost savings by reducing avoidable hospital admissions, readmissions, and emergency department visits. PGPs have flexibility in designing care management strategies to be successful under the demonstration. Several PGP participants initially focused on reducing avoidable admissions and readmissions among congestive heart failure patients, increasing influenza and pneumovax vaccine rates because of the potential for short-term payback, and improving transitions in care focusing on making sure newly discharged patients receive timely followup care. In addition, several PGP participants are focusing on a small number of very expensive patients, usually those who are hospitalized multiple times because these patients may show the largest effect from care management interventions by reducing avoidable readmissions.

PGP participants have been able to respond to the financial incentives under the demonstration, in many instances, through the enhancement and application of existing care management and information technology infrastructure and applying it to the

Medicare FFS population. PGP participants are relying mostly on in-house personnel, expertise, and enhancements to existing information technology infrastructure to implement their care management strategies, although some are partnering with, or have purchased systems from, outside vendors. In general, participants are strong proponents of provider-based care management because it builds off of their existing relationship with the patient and provides a consistent approach across the physician's practice making it easier to obtain physician buy-in.

PGP participants have indicated they view the demonstration quality measure methods used to assess and reward high performance and improvement as generally appropriate. However, most PGP participants commented that the additional resources required to collect data for medical record-based measures were more than expected due to the addition of the clinical record-based measures. But it is anticipated that the cost will decline somewhat in future years as the initial costs of developing systems and processes for collecting chart-based measures on a flow basis will not recur. The PGP participants' main strategies to improve their performance on the quality indicators are: (1) provider education and feedback including data profile reports comparing individual providers to their peers or other benchmarks, (2) better adherence to quality of care protocols on the part of both patients and physicians through care management interventions, and (3) implementation of standardized, evidence-based care models and protocols.

DISCUSSION

Medicare is exploring alternative approaches to improving the quality of care it pays for and controlling its costs. In the 1990s, managed care was a favored approach, but it has suffered setbacks in recent years (Robinson, 2001). More recently, P4P has been considered a promising approach (Bodenheimer et al., 2005). The PGP demonstration is Medicare's first physician P4P initiative. Unlike some other P4P initiatives, the PGP demonstration explicitly establishes incentives for cost efficiency as well as QI. It is a provider-based model that relies on the physician group as the organizational means to improve the quality and cost efficiency of care.

The PGP demonstration model changes provider payment, not the insurance arrangements of Medicare beneficiaries, who remain enrolled in the traditional FFS program with complete freedom of provider choice. Disruptions to providers are minimized by the maintenance of standard FFS Medicare payments to them. The innovation of the PGP demonstration model is that participating provider groups have the opportunity to earn additional performance payments for providing high quality and cost efficient care. They share the savings they create in the care of beneficiaries assigned to them with the Medicare Program, and retain more of the savings the higher their measured quality of care. The financial risk to providers is mitigated by the continuation of FFS payment, the use of provider-specific base costs as a starting point for measuring savings, and the lack of penalties for underperformance. Providers do face business risk for their investments in staff and systems to improve quality and generate savings, because they do not receive any upfront payments from Medicare. They may also forego some FFS revenues from rationalizing services provided, depending on whether additional demand for their services can be found to replace the foregone services.

The PGP demonstration payment model retains important FFS advantages as compared with capitation: for beneficiaries, freedom of provider choice, and for providers, no insurance risk and fewer incentives for stinting on services and avoiding the sickest patients. Theoretical analyses by health economists suggests that hybrid or mixed capitation/FFS models, such as the PGP demonstration model, may be preferable to either pure FFS or pure capitated payment (Ellis and McGuire, 1986).

In addition to establishing incentives to control volume of services, the PGP demonstration model takes advantage of the Medicare FFS program's market power to hold down the FFS rates paid to providers. Rather than relying on absolute cost control targets, the PGP demonstration model establishes yardstick competition (Schliefer, 1985) among providers by employing a local comparison group to judge cost-control performance. Expenditure growth performance targets are reasonable and feasible, because the baseline levels used for comparison have been achieved by other providers in the local market.

Like all payment innovations, the PGP demonstration faces some challenges. For example, it remains to be seen how much control a demonstration participant can exert over its assigned beneficiaries when they retain freedom of provider choice and have limited incentives to restrain their use of services.[7] In addition, the quality measures applied in this demonstration focus on ambulatory care for chronic diseases. Future efforts could reward additional types of quality, such as in acute care hospitalization and post-acute care episodes. Nevertheless, the PGP demonstration's attraction as a provider-based model combining elements of FFS and managed care make it worth testing as one of several possible approaches to Medicare payment reform. It will be important to monitor and evaluate the results of the PGP demonstration on several dimensions, including the characteristics of the participants and their patients; the implementation and operational experiences of participating PGPs; the impacts on Medicare Program expenditures, quality of care, providers, and beneficiaries. In addition, the generalizability of the demonstration model across various types and sizes of physician groups and geographic areas should also be assessed.

[7] As a provider payment model, the PGP demonstration could be combined with revised beneficiary cost sharing or other requirements to encourage beneficiaries to use services in a more cost efficient manner and to adhere to prescribed services that improve quality.

ACKNOWLEDGMENTS

The authors would like to thank John Pilotte and Heather Grimsley for their significant contributions to the design of the PGP demonstration, and for their insightful comments on the initial draft of this article.

REFERENCES

Baker, L.C.: Measuring Competition in Health Care Markets. *Health Services Research* 36(1, Pt. 2): 223-251, April 2001.

Bodenheimer, T., May, J.H., Berenson, R.A., et al.: *Can Money Buy Quality: Physician Response to Pay for Performance.* Center for Studying Health System Change. Issue Brief No. 102. December 2005. Internet address: http://www.hschange.com/CONTENT/807/ (Accessed 2007.)

Centers for Medicare & Medicaid Services: *Health Care Financing Review, Statistical Supplement, 2006.* Publication Number 03477. December 2006. Internet address: http://www.cms.hhs.gov/review/default.asp (Accessed 2007.)

Centers for Medicare & Medicaid Services: *CMS Physician Focus Quality Initiative: Chronic Disease and Prevention Measures.* November 2005. Internet address: http://cms.hhs.gov/PhysicianFocusedQualInits/downloads/PFQIClinicalMeasures200512.pdf (Accessed 2007.)

Ellis, R.P. and McGuire, T.G.: Provider Behavior Under Prospective Payment. Cost Sharing and Supply. *Journal of Health Economics* 5(2):129-151, June 1986.

Kautter, J., Pope, G.C., Trisolini, M., et al.: *Physician Group Practice Demonstration Bonus Methodology Specifications*. Final Report to the Centers for Medicare & Medicaid Services under Contract Number 500-00-0024(TO13). RTI International. Waltham, MA. December 2004.

Kautter, J., Pope, G.C., Trisolini, M., et al.: *Physician Group Evaluation First Evaluation Report*. Final Report to the Centers for Medicare & Medicaid Services under Contract Number 500-00-0024 (TO13). RTI International. Waltham, MA. June 2006.

Medicare Payment Advisory Commission: *Payment for Physician Services in the Medicare Program.* Testimony before the Subcommittee on Health Committee on Energy and Commerce, U.S. House of Representatives. May 5, 2004.

McCall, N., Pope, G.C., Adamache, K., et al.: *Research and Analytic Support for Implementing Performance Measurement in Medicare Fee for Service: First Annual Report.* Prepared for the Health Care Financing Administration under Contract Number 500-95-0058(TO2). Health Economics Research. Waltham, MA.1998.

National Committee for Quality Assurance: *HEDIS® 2006: Health Plan Employer Data & Information Set*. Washington, DC. 2005. Internet address: http://web.ncqa.org/tabid/292/Default.aspx (Accessed 2007.)

Olmsted, E., Pope, G.C., and Kautter, J.: *PGP Demonstration Report on Risk Adjustment*. Final Report to Centers for Medicare & Medicaid Services under Contract Number 500-00-0024(TO13). RTI International. Waltham, MA. 2006.

Pope, G.C., Kautter, J., Ellis, R.P., et al.: Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model. *Health Care Financing Review* 25(4):119-141, Summer 2004.

Pope, G.C., Trisolini, M., Kautter, J., et al.: *Physician Group Practice (PGP) Demonstration Design Report*. Final Report to the Centers for Medicare & Medicaid Services under Contract Number 500-95-0048(TO4). RTI International. Waltham, MA. October 2002.

Pope, G.C. and Chromy, J.: *Issues in Demonstrating Group-Specific Volume Performance Standards.* Prepared for the Health Care Financing Administration under Contract Number 500-95-0048(TO4). Health Economics Research, Inc. Waltham, MA. 1997.

Robinson, J.C.: The End of Managed Care. Journal *of the American Medical Association* 285 (20):2622-2688, May 2001.

Schliefer, T: A Theory of Yardstick Competition. *Rand Journal of Economics* 16(3):319-327, Autumn 1985.

Tompkins, C.P., Bhalotra, S., Glavin, M.P.V., et al.: *A Demonstration of Group-Specific Volume Performance Standards (GVPS)*. Prepared for the Health Care Financing Administration. Waltham, MA. Brandeis University. September 24, 1996.

Trisolini, M., Kautter, J., Pope, G.C., et al.: *Physician Group Practice Demonstration Quality Measurement Specifications, Version 2*. Report to the Centers for Medicare & Medicaid Services under Contract Number 500-00-0024(TO13). RTI International. Waltham, MA. July 2005.

Wallack, S. and Tompkins, C.: Realigning Incentives in Fee-For-Service Medicare. *Health Affairs* 22(4): 59-70, July/August 2003.

Reprint Requests: John Kautter, Ph.D., RTI International, 1440 Main Street, Suite 310, Waltham, MA 02451. E-mail: jkautter@rti.org