

Physician Group Practice (PGP) Demonstration Design Report

Prepared by:

Gregory C. Pope, M.S.
Michael Trisolini, Ph.D., M.B.A.
John Kautter, Ph.D.
Walter Adamache, Ph.D.

*Health Economics Research, Inc.
411 Waverley Oaks Road, Suite 330
Waltham, Massachusetts*

for:

Centers for Medicare and Medicaid Services
Baltimore, Maryland

October 2, 2002

Gregory C. Pope, M.S.
Project Director

Jerry Cromwell, Ph.D.
Scientific Reviewer

The research presented in this report was performed under Centers for Medicare and Medicaid Services (CMS) Contract No. 500-95-0048/Task Order #4, John Pilotte Project Officer. The statements contained in this report are solely those of the authors and no endorsement by CMS should be inferred or implied.

This Design Report was submitted by Health Economics Research, Inc. to the Centers for Medicare and Medicaid Services (CMS) as the proposed design for the Physician Group Practice Demonstration. CMS reserves the right to make changes to the proposed design and to set demonstration policy.

Table of Contents

	<u>Page</u>
EXECUTIVE SUMMARY	ES-1
I. Introduction	1
II. Demonstration Eligibility and Beneficiary Assignment	9
Recommendation 1: Eligibility of PGPs for the Demonstration	9
Recommendation 2: Assigning Beneficiaries to PGPs	16
III. Bonus Computation.....	23
Recommendation 3: Comparison Population and Expected Growth Rate	23
Recommendation 4: Sharing Rate	32
Recommendation 5: Bonus Payments, Settlements, and Withdrawal	39
Appendix to Recommendation 5: Equations Behind Simulations.....	56
Recommendation 6: Withholds.....	57
Recommendation 7: Required Number of Beneficiaries for Participating Physician Group Practices and Comparison Groups, and Thresholds for Bonus Payment.....	61
Recommendation 8: Rebasing Expenditures	69
Recommendation 9: Preliminary Expenditure Targets and Interim Performance Reporting	74
IV. Expenditure Definition and Adjustments	80
Recommendation 10: Expenditure Definition	80
Recommendation 11: Medicare Pass-Throughs and PPS Add-Ons	85
Recommendation 12: Cost Outliers	88
Recommendation 13: Health Status Casemix Adjustment	90
Recommendation 14: Adjusting Performance Targets for Effects of Medicare Payment Policy	100

Table of Contents (continued)

	<u>Page</u>
V. Quality Targets and Bonuses.....	103
Recommendation 15: General Approach for Quality Indicators	104
Recommendation 16: Selecting Specific Process and Outcome Quality Indicators	105
Recommendation 17: Data Sources for Measuring Quality Indicators	121
Recommendation 18: Quality Targets for PGPs.....	123
Recommendation 19: The Portion of the PGP Bonus Pool That Can be Earned Through Quality Performance	130
Recommendation 20: Calculating and Allocating Process and Outcome Quality Improvement Bonuses	134

References

Appendix A	Patient Assignment Steps
Appendix B	Steps in Defining the Comparison Population and Calculating the Target Growth Rate
Appendix C	Potential Medicare Claims-Based Process and Outcome Quality Indicators
Appendix D	Ambulatory Care Sensitive Conditions
Appendix E	Simulations of Adjusted Sharing Rate for Teaching Hospitals; Discussion of Timeliness of IME Payments
Appendix F	Effective Sample Size of Market Area (Comparison Group) Beneficiaries

Table of Tables and Figures

		<u>Page</u>
Figure ES-1	Recommended Distribution of Medicare Savings in the PGP Demonstration.....	ES-4
Table ES-1	PGP Bonus and Medicare Program Savings as a Percentage of Demonstration Medicare Savings.....	ES-5
Figure ES-2	PGP Demonstration Timeline.....	ES-11
Table ES-2	Example of PGP Demonstration Bonus Computation.....	ES-13
Table ES-3	Recommended Quality Performance Targets for Each Quality Indicator.....	ES-33
Figure I-1	Recommend Distribution of Medicare Savings in the PGP Demonstration.....	5
Table I-1	PGP Versus Medicare Program Payments as a Percentage of Demonstration Medicare Savings.....	7
Table 1-1	MGMA Membership by Group Size, 1999.....	12
Table 1-2	AMGA Membership by Group Size, 2001.....	12
Table 2-1	Evaluation and Management Services Included and Excluded in Beneficiary Assignment Criteria.....	18
Table 2-2	Number of Beneficiaries Assigned to Physician Group Practice (PGP), 1997.....	20
Table 4-1	Parameter Values for Sharing Rate Simulation.....	33
Table 4-2	Simulated Bonus Payments for Various Sharing Rates.....	37
Table 4-3	Simulations of PGP Change in Medicare Revenues for Various Sharing Rates.....	38
Figure 5-1	PGP Demonstration Timeline.....	41
Table 5-1	Example of PGP Demonstration Bonus Computation.....	43
Table 5-2	Simulations of Per Capita Bonus Payments Under Alternative Expenditure Growth Rate Scenarios.....	45
Table 5-3	Maximum Per Capita PGP Bonus Payments by Performance Year (PY) Under Alternative Expenditure Growth Scenarios (Summary of Table 5-2).....	50

Table of Tables and Figures (continued)

		<u>Page</u>
Table 6-1	Cash Flow by Bonus Pool and Withholding Rate	60
Table 7-1	Effect of PGP and Market Area Sample Sizes on the Probability of Paying a Bonus	63
Table 7-2	Effect of Alternative Bonus Thresholds on Probability of Paying a Bonus	67
Table 8-1	Effect of Annual Versus No Rebased on Per Capita PGP Bonus Pool and Medicare Program Savings	71
Table 9-1	Template for Interim Utilization Reporting (for a Physician Group Practice)	79
Table 10-1	Per Capita Medicare Payments and Growth Rates, 1993 and 1994, by Medical Service Category.....	82
Table 12-1	Percent Change in Mean Medicare Payments, 1993-1994 (Excluding High Cost Users).....	89
Table 12-2	Distribution of Annualized 1997 Medicare Expenditures	91
Table 13-1	Hypothetical Example of Concurrent DCG-HCC Model Relative Risk Score for a Beneficiary	94
Table 13-2	Hypothetical Example of Casemix Adjustment of Expenditure Growth and Medicare Savings Calculation	96
Table 13-3	Simulated Bonus Payments With and Without Casemix Adjustment	97
Table 18-1	Recommended Quality Performance Targets For Each Quality Indicator	131

EXECUTIVE SUMMARY

I. Introduction

The physician group practice (PGP) demonstration is a unique reimbursement mechanism through which providers are rewarded for coordinating and managing the overall health care needs of a non-enrolled, fee-for-service (FFS) patient population. It offers an opportunity to test whether a different financial incentive structure can improve service delivery and quality for Medicare patients and ultimately prove cost-effective.

The PGP demonstration superimposes new incentives on traditional FFS reimbursement that are more in line with capitation incentives. PGP organizations will have an incentive to reduce utilization for Medicare FFS patients. However, organizations that do not reduce utilization are not penalized under the PGP demonstration. The PGP demonstration includes explicit incentives for quality improvement. Performance on both process and outcome quality indicators, together with cost savings, will be used in the calculation of performance bonuses.

A legislative mandate for the PGP demonstration was included in the Medicare, Medicaid, and State Child Health Insurance Program Benefits Improvement and Protection Act (BIPA) of 2000. The general intent of the PGP demonstration can be summarized from BIPA 2000 as including five goals:

- 1) To test the use of incentives for health care groups, including physicians and other providers.
- 2) To encourage coordination of health care furnished under Medicare parts A and B.
- 3) To encourage investment in administrative structures and processes for efficient service delivery.

- 4) To reward physicians for improving health care processes and outcomes.
- 5) To focus on analysis of the efficiencies and advantages of furnishing health care in a group-practice setting as compared to other health care delivery systems.

In addition to their standard Medicare FFS reimbursement, which they will continue to receive, PGPs participating in the demonstration will be eligible to earn annual performance bonus payments. PGPs can earn bonus payments for both efficiency and quality performance. To earn bonus payments, a PGP must generate positive Medicare savings. If a PGP does not generate positive Medicare savings, no cost or quality bonuses are paid, and no savings accrue to the Medicare program. A summary of the steps involved in calculating annual Medicare savings is as follows:

- 1) For beneficiaries assigned to a PGP in the base year, the base year per capita expenditures are calculated.
- 2) The expenditure target for the performance year is calculated as follows:

$$\text{Target} = \text{Adjusted Base Year Per Capita Expenditures} \times (1 + \text{Expected Growth Rate})$$

Per capita expenditures in the base year are adjusted to account for differences in the casemix of beneficiaries assigned to the PGP in the performance year. The expected growth rate is defined as the growth rate in per capita expenditures in the PGP's local market area (comparison group) between the base and performance years, adjusted for casemix change.

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

$$\text{Medicare Savings} = (\text{Target} - \text{Performance Year Per Capita Expenditures}) \times \text{FTE Beneficiaries.}$$

The PGP demonstration is envisioned as a three year demonstration. The concepts and steps involved in multi-year bonus computations are discussed below in the specific design recommendations.

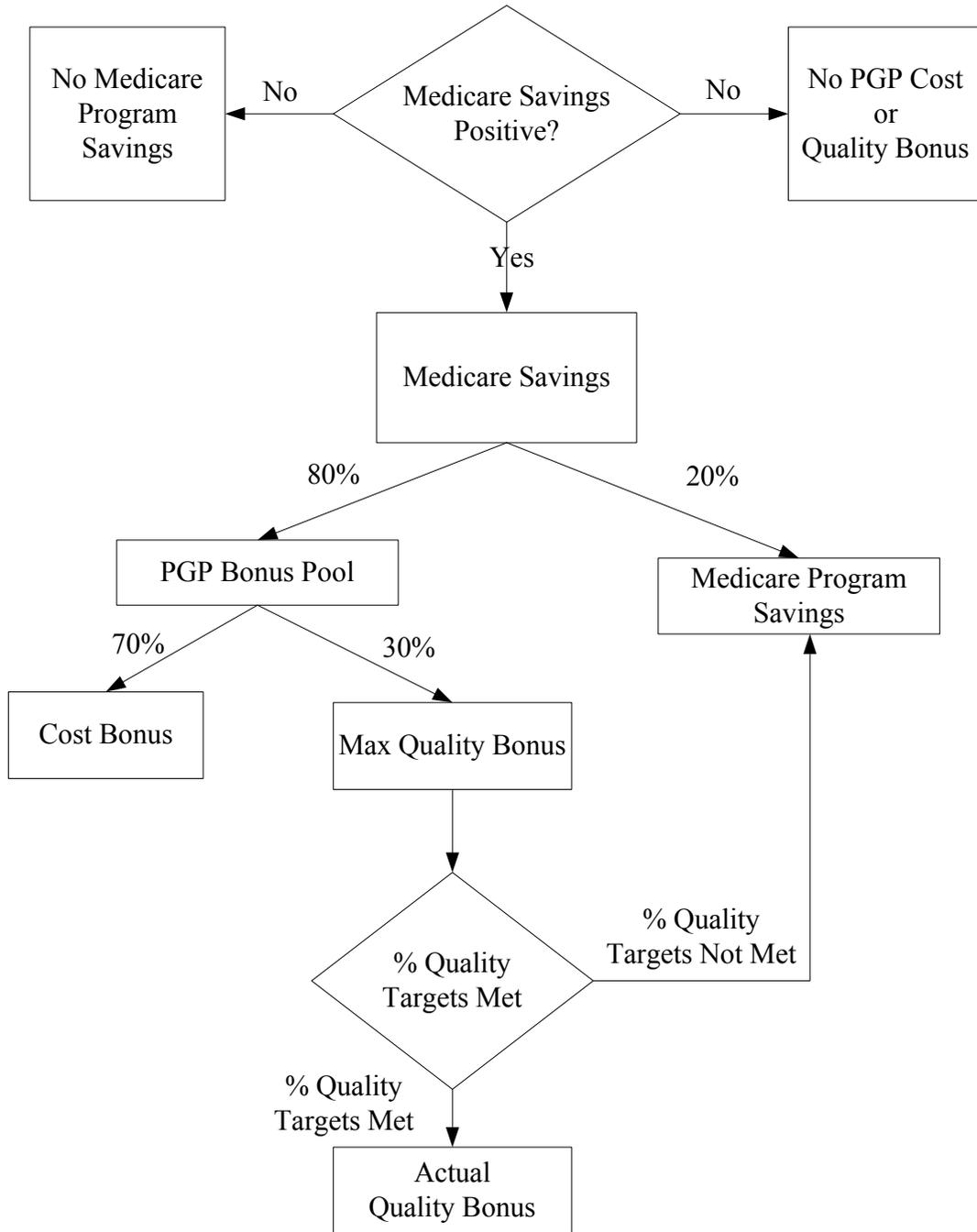
The distribution of Medicare savings in the PGP demonstration is shown in Figure ES-1. As shown in the Figure, if a PGP does not generate positive Medicare savings, no cost or quality bonuses are paid, and no savings accrue to the Medicare program.¹ For PGPs that do generate positive Medicare savings, the recommended sharing rate equals 80%, which means that 80% of Medicare savings is allocated to the PGP bonus pool, and the remaining 20% is savings for the Medicare program. The PGP automatically receives 70% of the bonus pool as a “cost bonus,” but must earn the remaining portion of the PGP bonus pool by providing high quality care. The maximum quality bonus the PGP can earn is the remaining 30% of the PGP bonus pool. The actual PGP quality bonus equals the maximum quality bonus multiplied by the percentage of quality targets met by the PGP. The remainder of the maximum quality bonus, if any, is additional savings for the Medicare program.

For PGPs that generate positive Medicare savings, Table ES-1 shows the maximum and minimum payments to PGPs and the Medicare program as a percentage of Medicare savings. PGPs will receive a minimum 56% of Medicare savings through their cost bonus (80% sharing rate x 70% cost bonus percentage = 56%). The maximum quality bonus is 24% of Medicare savings if all quality targets are met (80% sharing rate

¹ No Medicare savings are considered to be generated unless they exceed a threshold of 2% of target expenditures. See Recommendation 7.

Figure ES-1

Recommended Distribution of Medicare Savings in the PGP Demonstration



SOURCE: Health Economics Research, Inc.

Table ES-1
PGP Bonus and Medicare Program Savings
as a Percentage of Demonstration Medicare Savings

	<u>All Quality Targets Met</u>	<u>No Quality Targets Met</u>
PGP Bonus	80%	56%
Cost	56	56
Quality	24	0
Medicare Program Savings	20	44

NOTE: Assumes positive Medicare savings. Otherwise no cost or quality bonuses are paid, and no savings accrues to the Medicare Program.

SOURCE: Health Economics Research, Inc.

x 30% maximum quality bonus percentage x 100% quality targets met percentage = 24%). The minimum quality bonus is 0% of Medicare savings if no quality targets are met (80% sharing rate x 30% maximum quality bonus percentage x 0% quality targets met percentage = 0%). Thus, PGPs will receive 80% of the Medicare savings when all quality targets are met, versus 56% when no quality targets are met. Conversely, the Medicare program receives 44% of Medicare savings when no quality targets are met, versus only 20% when all quality targets are met.

This Design Report includes 20 specific recommendations regarding the design of the PGP demonstration. They involve key issues which must be addressed in order to implement the demonstration effectively. The recommendations are as follows:

II. Demonstration Eligibility and Beneficiary Assignment

Recommendation 1: Eligibility of PGPs for the Demonstration

The PGP demonstration should include several different types of PGPs in order to test its new incentives in a range of organizational and clinical environments. In particular, HER recommends that the demonstration be open to:

- Stand-alone PGPs as well as those affiliated with hospitals, other provider organizations, or integrated delivery systems (IDSs); and
- PGPs with national reputations as tertiary referral centers as well as smaller, regional providers.

However, HER also recommends that four more restrictive eligibility criteria also be applied. They relate to the capability of participating PGPs to respond effectively to the demonstration's new incentives. In particular, the demonstration should be restricted to:

- Large PGPs, with at least 200 physicians and at least 15,000 to 20,000 Medicare FFS beneficiaries assigned;
- Multispecialty PGPs;
- PGPs with relatively well developed information, clinical, and management systems; and
- Contracting entities which are either stand-alone PGPs or distinct organizational units defined as physician groups within larger IDSs or other multi-provider organizations.

Recommendation 2: Assigning Beneficiaries to Physician Group Practices

- Beneficiaries who receive at least one Evaluation and Management (E&M) service from a participating PGP are eligible for assignment to the PGP.
- Beneficiaries with any Medicare+Choice enrollment are not eligible for assignment.
- Beneficiaries who receive more E&M services (measured by Medicare payments) from a participating PGP than from any other physician practice (group or solo) will be assigned to the PGP.
- Certain E&M services will be excluded from assignment, including consultations, emergency department visits, and critical care services.
- Beneficiaries will be assigned to at most one participating PGP.

Since the PGP demonstration is a fee-for-service innovation, there is no enrollment process whereby beneficiaries accept or reject participation. Thus beneficiaries need to be “assigned” to PGPs based on their utilization of Medicare-covered services. The intention of the demonstration is to give each participating PGP an incentive to manage the health care of the beneficiaries assigned to it. The ability of a physician practice (group or solo) to manage the health care of a beneficiary depends on its control over the beneficiary’s utilization of services. Because of this, a participating PGP providing the largest share of E&M services is, we believe, in the best position to manage the health care of the beneficiary. Unique assignment to PGPs prevents CMS from paying bonuses more than once when multiple PGPs serve overlapping Medicare patient populations. Also, because assignment is based only on E&M service codes, assignment does not depend on physician specialty.

Certain E&M services do not reflect the PGP’s ability to manage the health care of the beneficiaries assigned to it, and thus should not be used for assignment. We

recommend that the following E&M services be excluded from use in assignment: consultations, emergency department visits, and critical care services.

We are still considering options for several scenarios. First, a beneficiary may receive an equal amount of E&M services from two participating PGPs during a year, and no E&M services from any other physician practice (group or solo). Under this scenario, the beneficiary would not be assigned to either PGP under the recommended assignment criteria. Second, and potentially more important, is the possibility that a PGP provides the majority of Medicare-covered services to a beneficiary during a year, and yet the beneficiary is not assigned to the PGP under the recommended assignment criteria. Assignment criteria may be modified after further analysis of these situations.

III. Bonus Computation

Recommendation 3: Comparison Population and Expected Growth Rate

- The comparison population for a participating PGP will be fee-for-service beneficiaries residing in its market area who are not assigned to the PGP. The expected growth rate will be the change in Medicare per capita expenditures for comparison group beneficiaries from the base to the performance year.

The comparison population for a participating PGP is fee-for-service Medicare beneficiaries residing in the PGP's market area that is not assigned to the PGP. The PGP's market area will be defined as counties in which 1% or more of the beneficiaries assigned to the PGP reside. The market area will be defined for both base and performance years, and may differ between the two years to reflect changes in the PGP's service area (although in general we expect PGP service areas to be stable over time).

The PGP's expected expenditure growth rate will be the change in market area per capita expenditures from the base to the performance year. Market area per capita expenditures will be defined as weighted average county per capita expenditures of market area counties. The weights will be the share of participating PGP beneficiaries residing in each market area county. The comparison group of each participating PGP will be required to have at least 15,000 assigned beneficiaries.

Recommendation 4: Sharing Rate

- A sharing rate of 80 percent will be used to determine bonus payments.

The sharing rate is the maximum proportion of the Medicare savings generated by a PGP that can be paid to the PGP as a bonus. The sharing rate needs to be high enough to give PGPs sufficient incentive to participate in the demonstration, but low enough so that the Medicare program shares significantly in any savings. Based on HER's simulations of potential PGP bonuses under the demonstration, a sharing rate of 80% percent is recommended. With this sharing rate, the PGP earns up to 80% percent of the Medicare savings it generates (depending on its performance with regard to the quality of care targets), and a minimum of 20% of Medicare savings accrues to the government.

Recommendation 5: Bonus Payments, Settlements, and Withdrawal

- Bonuses may be earned by participating PGPs in performance years in which the organization generates Medicare Savings.
- Medicare Losses accrue to participating PGPs in performance years in which PGP expenditures exceed their Target.

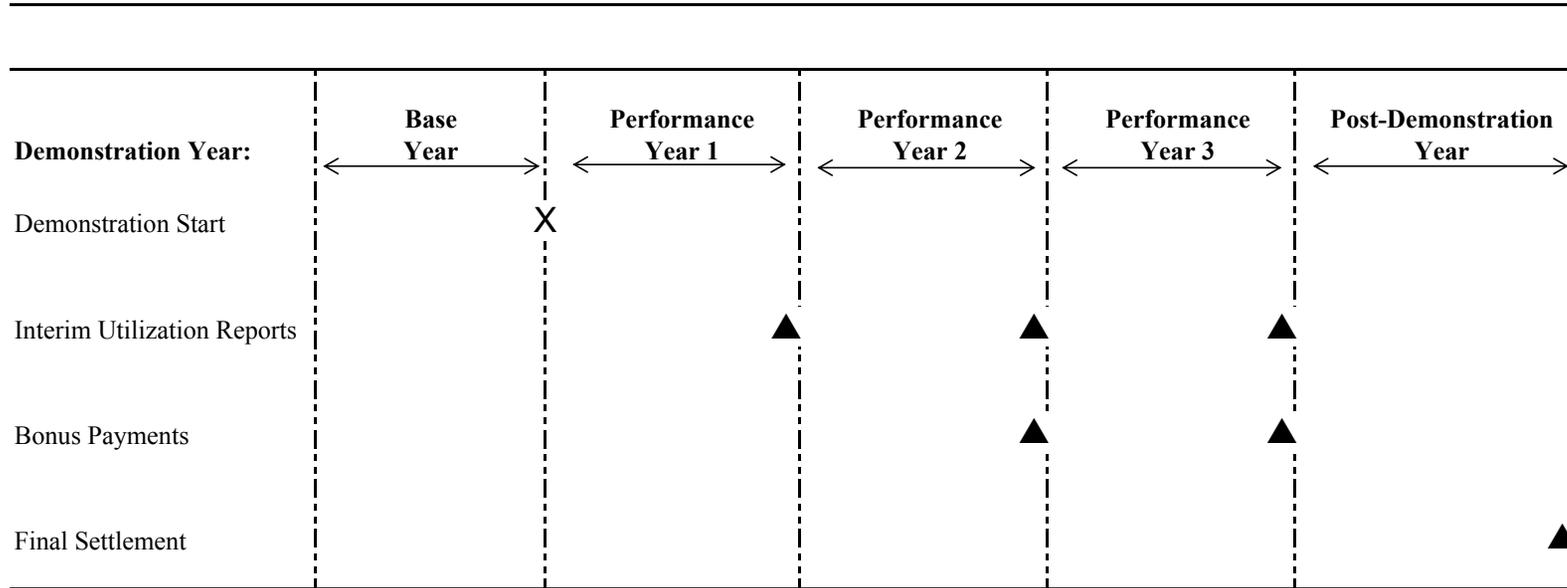
- A PGP's accrued Medicare Loss carried forward from the prior performance year, if any, is deducted from its Medicare Savings before bonuses are determined.
- The annual PGP bonus, if any, will be paid at annual settlement, with a portion withheld until final demonstration settlement contingent on future performance.
- The maximum bonus that can be earned by a PGP in a year (bonus payments plus withheld amounts) is limited to 15% of target Medicare expenditures for beneficiaries assigned to that organization in that year.
- At final demonstration settlement, CMS will remit withheld bonus amounts to the PGP. Accrued losses will be deducted from the amount returned by CMS to the PGP. Even if accrued losses exceed withheld bonuses, at most the PGP will forfeit withheld bonuses at final settlement.
- If a participating PGP withdraws from the demonstration before its completion, it will forfeit all withheld bonus payments.

Figure ES-2 shows the demonstration schedule. The demonstration will begin with performance year one (PY1). An interim report on utilization in PY1 versus the base year will be provided to the participating PGP at the end of PY1 (see Recommendation 9). Medicare claims data for PY1 will be considered complete six months after the end of the year. Claims data for the PGP and its comparison group will be obtained from Medicare datafiles and processed by the demonstration technical support contractor over the next 6 months. An annual bonus settlement cycle will occur approximately one year after the end of PY1, and then one year after PY2 and one year after PY3. Final demonstration settlement will occur approximately one year after the end of PY3, simultaneously with the PY3 annual bonus settlement.

Bonus payments may be made to a PGP if it generates Medicare savings in a performance year. A "Medicare Loss" will accrue in any performance year in which

Figure ES-2

PGP Demonstration Timeline



SOURCE: Health Economics Research, Inc.

PGP per capita expenditures exceed its Target.² The PGP's bonus pool in a performance year is based on the PGP's Medicare Savings or Loss for that year combined with its accrued Medicare Losses, if any, carried forward from the prior performance year. If this combined amount is greater than zero, a portion of the amount will be paid to the PGP and the remainder will be withheld contingent on future performance. (Withheld amounts will be returned to the PGP as part of final settlement.) To avoid incentives for excessive cost cutting, the maximum bonus that can be earned by a PGP in a year (bonus payments plus withheld amounts) is limited to 15% of target Medicare reimbursements for assigned beneficiaries. Participating organizations would simply forgo bonuses in excess of this limit, i.e., bonuses in excess of the 15% limit will not be paid, withheld, or accrued for payment in future years.

An example, as shown in Table ES-2, will help clarify these concepts and steps. We simplify the example by assuming that a participating organization meets all quality targets each year of the demonstration so that the PGP earns all of its bonus pool. Suppose a participating organization generates positive Medicare Savings of \$3,900,000 in performance year 1 (PY1), a Medicare Loss of -\$6,030,000 in PY2, and Medicare savings of \$4,350,000 in PY3. After applying the Sharing Ratio of 80%, the PGP's bonus pool for PY1 is \$3,120,000. Assuming a withholding rate of 25% (see Recommendation 6), \$2,340,000 is paid to the PGP as the PY1 bonus and \$780,000 is withheld for final settlement. In PY2, Medicare Savings are negative. A Loss of -\$6,030,000 is accrued, and is carried forward to PY3. In PY3, Medicare savings are

² Small annual Medicare Savings or Losses, that may be due to chance, will be considered to be zero. See Recommendation 7.

Table ES-2

Example of PGP Demonstration Bonus Computation

	<u>Performance</u> <u>Year 1</u>	<u>Performance</u> <u>Year 2</u>	<u>Performance</u> <u>Year 3</u>	<u>Final</u> <u>Settlement</u>
Annual Medicare Savings or Loss	\$3,900,000	-\$6,030,000	\$4,350,000	--
Accrued Medicare Loss from Previous Year	0	0	-6,030,000	-1,680,000
Sum ¹	3,900,000	-6,030,000	-1,680,000	--
Annual Bonus Pool ²	3,120,000	0	0	--
Bonus Paid ³	2,340,000	0	0	--
Bonus Withheld Until Final Settlement ⁴	780,000	0	0	780,000
Final Settlement Amount ⁵	--	--	--	0

¹Sum of Annual Medicare Savings/Loss and Accrued Medicare Loss from previous year.
If negative, carried forward to next year.

²80% (the sharing ratio) of the sum of annual accrued Medicare Savings/Losses if greater than zero, zero otherwise.

³75% (100% - the withholding %) of the bonus pool.

⁴25% (the withhold %) of the bonus pool.

⁵Sum of 80% of accrued Medicare Loss from Year 3, if any, and bonus withheld until final settlement.
In this example, this sum equal -564,000. However, the participating PGP never loses more than
the withheld bonus amounts at final settlement. So the final settlement amount is zero in this example.

NOTE: Assumes that all quality targets are met each year, a sharing ratio of 80%, and a bonus
withholding rate of 25%.

SOURCE: Health Economics Research, Inc.

\$4,350,000. But when the accrued Loss of -\$6,030,000 carried forward from the previous year is charged against the Annual Savings, the resulting amount for PY3 is -\$1,680,000 and no PY3 bonus is paid.

Final demonstration settlement will occur approximately one year after the end of PY3. At settlement, the withheld annual bonus payments will be debited by 80% of the Accrued Loss, if any, at the end of the demonstration.³ Continuing the example of the preceding paragraph, withheld bonuses total \$780,000. Eighty percent of the Accrued Loss at the end of the demonstration is -\$1,344,000 (80% of -\$1,680,000). The sum is -\$564,000. Therefore, the PGP forfeits the \$780,000 in withheld bonuses, i.e., CMS does not return the withheld bonuses at demonstration settlement. The participating organization is never required to return to CMS at settlement any paid bonuses it has received during the demonstration.

The PGP demonstration is intended to measure longer-run, permanent changes in PGP cost and quality behavior, not transitory year-to-year fluctuations. For this reason, if a participating organization withdraws from the demonstration prior to its full, three-year completion, it will forfeit all withheld bonus payments. This policy avoids incentives for the PGP to drop out when it earns large bonuses in PY1 and PY2, but foresees a decline in its performance in PY3 (or simply does not want to chance a decline in performance in PY3) that would be charged against its earlier withheld bonuses.

Recommendation 6: Bonus Withholds

- HER recommends a 25 percent withholding rate on bonus payments.

³ The Accrued Loss is adjusted by the sharing ratio, 80%. Just as Medicare Savings are shared between the PGP and CMS, so too losses are shared.

- At the end of the demonstration, cumulative withheld amounts in excess of accrued losses will be returned to the PGP.

Over time, a participating PGP might earn bonuses in some years and accrue losses in other years. The question is whether bonuses should be paid in the year they are earned, or whether some portion should be withheld to offset future losses.

The main argument against withholds is that participating PGPs might need the demonstration bonus funds to offset the lower Medicare FFS revenues that generated the bonus. This may be especially relevant for PGPs with affiliated hospitals that generated their bonus by reducing hospital admissions, which generally involve larger revenues than other types of medical services.

An argument in favor of withholding is that it is administratively more feasible than CMS having to recover payments after the demonstration is over from participating PGPs. This is especially true if the magnitude of payments that need to be recovered are relatively small. Withholds also provide a measure of protection to the Medicare program in the event that a participating PGP that had received bonus payments decides to leave the demonstration prior to its scheduled end. That is, if it is difficult for Medicare to recover bonus payments, withholds ensure that Medicare can recover some of the bonuses earned by a withdrawing PGP.

The goal is to set a withholding rate that is low enough so that total Medicare payments are sufficient to induce organizations to participate in the demonstration, while high enough to provide the Medicare program with financial protection. Our simulations show that a withholding rate of 25 percent strikes a good balance between these two

goals. Higher withholding rates (e.g., 50%) tend to turn total Medicare payments to participating organizations negative (before withholds are returned at final settlement), while lower withholding rates provide limited financial protection to the Medicare program.

Recommendation 7: Thresholds for Bonus Payment

- In any performance year, if participating PGP Medicare Savings or Losses are within plus or minus 2% of target expenditures, then Medicare Savings are considered to be zero. That is, bonus payments are subject to a Medicare savings *threshold* of 2% of target expenditures. When Medicare savings exceed 2%, bonus payments may be made to PGPs, and when Medicare losses exceed 2% they will be accrued.

A bonus threshold avoids paying a bonus for small differences in actual versus target expenditures that could be due to chance. Choosing an appropriate bonus threshold involves trading off the probabilities of paying deserved bonuses versus not paying undeserved bonuses. Based on our simulations, we recommend a bonus threshold of 2.0%. This means that a bonus would not be paid unless the difference in actual and target expenditures exceeds 2.0%. If the threshold is exceeded, however, the full bonus is paid, not just the amount above the 2% threshold. Similarly, Medicare losses will only be accrued when the losses exceed 2%. If the loss threshold is exceeded, however, the full loss is accrued, not just the amount in the excess of the 2% threshold.

Recommendation 8: Rebasing Expenditures

- Participating PGP per capita expenditures will not be rebased during the demonstration.

Rebasing means changing the base year for the PGP bonus calculation. If the PGP demonstration is rebased annually, then the performance of the participating PGP is judged independently on an annual basis. If the demonstration is never rebased, the performance of the PGP is judged cumulatively over time. With no rebasing, the potential bonuses for participating PGPs are larger. With annual rebasing, more of the demonstration cost savings will be captured by the Medicare program. Over the relatively short period of the demonstration (3 years), HER recommends no rebasing.

Recommendation 9: Preliminary Expenditure Targets and Interim Performance Reporting

- Preliminary expenditure targets will not be reported in the demonstration, but interim utilization reports will be provided to participating PGPs.

HER recommends that CMS does not announce preliminary expenditure targets. Preliminary targets are not likely to be available soon enough to be useful, they are likely often to be inaccurate, and they are administratively burdensome to compile. Given data availability and processing lags, final PGP performance reporting (Medicare savings, bonuses) for a year will not be available until 9 to 12 months following the end of the year. Earlier interim reporting is desirable to give participating PGPs more timely feedback about their performance. HER recommends generating interim utilization reports for participating PGPs comparing hospital admission rates between the first six months of a performance year and the first six months of the prior year, for the beneficiary population assigned to the PGP in the prior year. These reports will be available approximately at the end of each performance year.

IV. Expenditure Definition and Adjustments

Recommendation 10: Expenditure Definition

- Use all Medicare Part A and Part B expenditures to calculate per capita expenditures for the demonstration.

BIPA 2000 requires that the PGP demonstration include "a base expenditure amount, equal to the average total payments under Parts A and B for patients served by the health care group on a fee-for-service basis in a base period determined by the Secretary". Since the primary goal of the PGP demonstration is to reduce the growth rate of Medicare fee-for-service expenditures, setting a comprehensive target gives the PGP more flexibility to focus on the largest sources of inefficiency. It encourages PGPs to take a comprehensive view when managing care and developing ways to better coordinate care, and avoids incentives for inefficient substitution of services to earn bonuses. All Medicare Part A and Part B expenditures are recommended for calculation of per capita expenditures in the demonstration.

Recommendation 11: Medicare Pass-Throughs and Add-Ons

- PPS pass-throughs and add-ons will be included when calculating per capita expenditures for the demonstration.

Two special categories of payments that hospitals receive from Medicare are "pass-throughs" and "add-ons". Pass-throughs include payments for direct graduate medical education, bad debt, costs of acquiring organs for transplants, and the nearly phased-out pass-through capital costs. PPS add-ons include payments to qualifying hospitals serving a disproportionate share of poor patients (DSH) and indirect graduate

medical education (IME) payments to hospitals with residency programs. Some participating PGPs may include an affiliated teaching hospital. If the PGP generates Medicare savings by reducing admissions at the teaching hospital, the organization will forgo PPS passthrough and add-on payments for the avoided admissions. The question is whether some special adjustment should be made in the PGP demonstration to reimburse participating organizations with teaching hospitals for forgone PPS passthrough and add-on payments.

HER's recommendation is to treat pass-throughs and PPS add-ons like any other expenditure and include them when calculating per capita expenditures for the demonstration. The participating PGP with a teaching hospital can get up to 80 percent of the foregone add-ons and passthroughs back through the demonstration bonus payment (exact amount depending on quality performance). In addition, the organization should enjoy cost savings associated with the forgone admissions. If these cost savings are at least 20 percent of the foregone add-on/passthrough revenues, the participating organization can come out ahead on add-ons and passthroughs under the demonstration.⁴ For these reasons, HER believes that organizations, including teaching hospitals, have sufficient incentive to participate in the demonstration without any special compensation for foregone PPS add-on and passthrough payments. Medicare payment policy (including BBA treatment of teaching pass-throughs in Medicare+Choice) has consistently treated PPS add-ons (IME/DSH) as variable costs paid on a per admission basis. HER's recommendation is consistent with this well-established Medicare policy.

⁴ Savings associated with forgone admissions may be greater for some pass-throughs/add-ons than others. Direct medical education costs, for example, may vary less with the number of admissions than IME/DSH costs.

It would be an extraordinary exception, probably not generalizable beyond the demonstration setting, to guarantee participants with a teaching hospital fixed amounts of IME and DSH payments independent of their number of admissions.

Recommendation 12: Cost Outliers

- For purposes of demonstration bonus calculations, annualized expenditures for any beneficiary assigned to a participating PGP or its comparison group will be truncated (capped) at \$100,000.

Random variability of expenditure growth rates for PGP demonstration participants or their comparison populations may reward PGPs even when they have not altered their behavior, and conversely, might lead to a lack of bonuses even when participants are reducing services per beneficiary. A small group of extremely costly beneficiaries assigned to a PGP could significantly change a PGP's per capita expenditures and hence its bonus. Thus, for each beneficiary assigned to a PGP or comparison group, HER recommends that annualized expenditures be truncated (capped) at \$100,000. In 1997, more than 99% of Medicare FFS beneficiaries had annualized expenditures less than \$100,000. A primary objective of the PGP demonstration is to give PGPs an incentive to manage the health care of the high cost patients assigned to them. Capping expenditures at a level lower than \$100,000 runs the risk of jeopardizing this objective.

Recommendation 13: Health Status Casemix Adjustment

- PGP and market area per capita expenditures will be adjusted for casemix changes using the Diagnostic Cost Groups-Hierarchical Condition Categories concurrent risk adjustment model.

When making comparisons between participating PGP and comparison group expenditure growth rates, health status casemix is held constant. The per capita expenditures of both participating PGPs and their comparison groups are adjusted for changes in casemix using the concurrent Diagnostic Cost Groups, Hierarchical Condition Categories (DCG-HCC) model. This model uses diagnoses on Medicare claims to predict the expected average expenditures of a population based on its health status. The DCG-HCC model is part of the same family of Diagnostic Cost Group models as the Principal Inpatient DCG (PIP-DCG) model that is currently used for risk adjustment of capitation payments to Medicare+Choice plans. However, it differs in two key respects from the PIP-DCG model used in M+C payment.

First, since ambulatory diagnoses are available from Medicare fee-for-service claims, the DCG-HCC model utilizes them. The PIP-DCG model utilizes only inpatient encounter claims as they are all that is available for M+C plans. Second, the PIP-DCG model is prospective, meaning that it forecasts expenditures in the following year. The model we recommend for this demonstration is concurrent, and explains expenditures in the current year. A concurrent model is more appropriate for current year casemix adjustment with a non-enrolled population that varies from year to year. The concurrent DCG-HCC model will need to be customized for this demonstration.

Casemix adjustments will be made by adjusting base year PGP and market area comparison group casemix to performance year casemix. This will be done by

multiplying base year expenditures by the ratio of the performance year to the base year casemix index.

Recommendation 14: Adjusting Performance Targets for Effects of Medicare Payment Policy

- Medicare payment policy changes will be reflected in the expenditures of both the participating PGP and its comparison group. Hence, Medicare savings and PGP bonuses under the demonstration will not be sensitive to changes in Medicare payment policies. No adjustments for changes in payment policy are recommended.

The PGP demonstration compares the per capita expenditure growth rate of a participating organization to an expected growth rate defined by its local market area comparison group. Except in unusual circumstances, changes in Medicare payment policy should similarly affect the expenditures of both the participating PGP and its comparison group. When the expenditure growth rates of the two are compared, effects of changes in payment policy should "cancel out". Hence, adjustments for the differential effects of changes in payment policy are not necessary.

V. Quality Targets and Bonuses

Recommendation 15: General Approach for Quality Indicators

- HER recommends the demonstration include 8 process and outcome quality indicators.
- CMS will work with participating PGPs to identify the final set of indicators.
- CMS will retain the right to make the final selection of indicators.

We believe a relatively small number of indicators (less than 10) will have several advantages for the demonstration. First, it reduces the complexity and cost of collecting data and calculating performance comparisons and incentive payments. Second, a smaller number of indicators can be focused on high volume clinical conditions which affect a large percentage of the Medicare population, or on conditions known to be associated with frequent or significant quality problems. Third, focusing on higher volume conditions will provide larger sample sizes to improve the statistical reliability of comparisons between quality targets and actual PGP performance.

The number of indicators could be increased in the future if this new PGP reimbursement system is implemented more widely by Medicare after the demonstration is completed. However, HER recommends that the short-run goal for the demonstration should be to focus on “pilot testing” a smaller set of indicators.

Recommendation 16: Selecting Specific Process and Outcome Quality Indicators

- The major focus of the demonstration will be on measuring process indicators of quality. They are the indicators most easily measured and most relevant to the medical care operations of PGPs.
- Selected outcome indicators will also be included.

Quality of care can be measured and analyzed using several different concepts. The basic choices are: (i) structure (or inputs); (ii) process; and (iii) outcomes (including patient satisfaction). The BIPA 2000 legislation mandates a focus on process and outcome indicators for the PGP demonstration.

Process indicators are more commonly used in health care quality improvement programs since they are more directly under the control of medical providers. Moreover, process indicators generally provide for annual or more frequent data collection, which will facilitate the annual monitoring efforts required for the PGP demonstration. Outcome indicators are sometimes used in quality improvement efforts, but they are often less practical to measure and analyze. Nonetheless, outcomes are what beneficiaries care about in the end, so we recommend including selected outcome indicators, focusing on those where the measurement issues are less of a concern.

We anticipate the specific process and outcome indicators used in the demonstration will be selected from among the following indicators recommended by HER:

Recommended Process Indicators

- Eye examination every 2 years for beneficiaries with diabetes.
- Hemoglobin A1c test every year for beneficiaries with diabetes.
- Lipid profile test every 2 years for beneficiaries with diabetes.
- Mammogram every 2 years for female beneficiaries between the ages of 52-69.
- Both chest radiograph and electrocardiogram ≤ 3 months after initial diagnosis of congestive heart failure (CHF).
- Left ventricular ejection fraction (LVEF) test conducted during the current year for beneficiaries hospitalized with a principal diagnosis of CHF during the current year.
- One physician visit every 6 months for beneficiaries with any one of the following 4 chronic disease diagnoses: chronic stable angina; chronic obstructive pulmonary disease (COPD); CHF; or diabetes.
- Annual influenza vaccinations for all beneficiaries age 65 or older.
- Pneumonia vaccination status for all beneficiaries age 65 or older (ever had a pneumonia vaccination).

Recommended Outcome Indicators

- Number of hospital admissions per year for three high volume ambulatory care sensitive conditions (ACSCs) per 1000 Medicare beneficiaries. The conditions include asthma/COPD, CHF, and bacterial pneumonia.
- Patient satisfaction maintained above a lower bound level (specific survey instrument and target level to be determined).

We considered four factors in selecting specific quality indicators to recommend for the PGP demonstration. First, focusing on indicators already in widespread use in national quality improvement efforts, such as the Health Plan Employer Data & Information Set (HEDIS) and Medicare's Health Care Quality Improvement Program (HCQIP), to ensure they will have credibility with participating PGPs. Second, including indicators that have been developed for – or applied to – the Medicare population, and published in peer-reviewed literature. Third, including indicators that represent a range of different types of medical care interventions. Fourth, focusing on indicators that apply to high volume diseases or clinical conditions to ensure sufficient sample sizes for reliably evaluating the performance of PGPs.

Recommendation 17: Data Sources for Measuring Quality Indicators

- Administrative billing ("claims") data will be the primary data source for measuring quality indicators.
- Patient survey data will also be collected to measure patient satisfaction and influenza and pneumonia vaccination rates.

Claims are low cost and take advantage of the central role for claims data in the PGP demonstration for calculating cost targets, cost performance comparisons, and

Medicare Savings. Patient surveys enable collection of data not available in claims, including patient satisfaction and influenza and pneumonia vaccination rates.

Other options for data sources include medical record abstracts, to provide more detailed clinical data, and additional patient surveys to collect data on health-related quality of life (HRQOL) and other self-reported information from beneficiaries. Unfortunately, medical record abstracts are costly. We believe they are unnecessary for this demonstration given the limited number of indicators to be measured and the availability of sufficient numbers of validated claims-based indicators. Some types of clinical information are available in claims data (e.g., diagnoses, utilization of specific tests and services), and we believe these will be sufficient. Eight of the eleven specific quality indicators recommended above have been analyzed in previous studies using Medicare claims data, and a number of other claims-based indicators have been developed and tested in recent years.

By recommending surveys focused on patient satisfaction we include some self-reported data for measuring the patient's own experience of health care. Collecting additional survey data on HRQOL would provide another type of self-reported information, but would add significantly to the length of the survey and add to the expense and complexity of the demonstration. With the limited number of indicators to be measured for this demonstration we believe the patient satisfaction data will be sufficient.

Recommendation 18: Quality Targets for PGPs

- The demonstration will include two types of quality targets a PGP can meet in order to earn the quality portion of the PGP Bonus Pool:
 1. Achieving a pre-determined threshold level for a quality indicator.
 2. Demonstrating substantial quality improvement over time.
- Only a threshold target is recommended for patient satisfaction, since that indicator is primarily intended to provide a lower bound for detecting undertreatment.
- *Either* threshold or quality improvement targets could be met to earn a quality bonus for all of the other indicators.

An advantage of pre-determined, fixed thresholds is that they reward physician groups for quality achievements involving evidence-based goals. A disadvantage of the threshold approach is that it may discourage some PGPs from applying for the demonstration if the thresholds are set at very high levels (e.g., requiring 95 percent of diabetics to receive hemoglobin A1c tests each year), which might be hard for any PGP to achieve.

An advantage of setting targets based on substantial quality improvement over time is that they are consistent with the goals of continuous quality improvement (CQI) or total quality management (TQM) programs, which underlie many of the recent efforts toward quality improvement in the health sector (IOM, 2001a). This perspective recognizes that achieving “ideal” thresholds is often very difficult in actual medical care practice settings with real-world patient populations. Instead of focusing on penalizing physicians or PGPs for failing to reach ambitious thresholds that may be somewhat beyond their control (at least in the short run), this approach rewards the realistic quality improvements that are possible in the short run (Chassin, 1996).

However, we also note that a disadvantage of quality targets focused on demonstrating improvement over time is that they may penalize groups that have already achieved high levels of quality. Those PGPs may find it much harder to show improvement than other groups starting from a base of much lower quality. As a result, our recommended approach allows PGPs to earn quality bonuses in either of the two ways, avoiding the problems involved in using one method exclusively.

We recommend the following protocols be used to determine if a PGP has met either the fixed threshold or the substantial improvement over time criterion for each indicator:

Pre-determined Threshold Targets

- The fixed threshold level of quality will be set at 75 percent of a PGP's eligible beneficiaries receiving the care specified by the indicator for six of the eight claims-based indicators (all except the indicators for ACSCs and visits every 6 months for beneficiaries with the four specified chronic conditions), and for the two survey-based vaccination measures.

For example, if 75 percent of a PGP's diabetics had hemoglobin A1c tests, it would have met the quality target. Further improvement would not be necessary to continue to earn the quality bonus for that indicator. The threshold is set somewhat below 100 percent in recognition that perfect performance is usually not achievable, due to: (a) limited PGP resources; (b) imperfect internal PGP clinical and management systems; (c) some degree of patient non-compliance with physicians' recommendations; and (d) other factors. The 75 percent threshold is, however, significantly above the national average rates for each of these indicators. Studies by Asch *et al.* (2000) and

Jencks *et al.* (2000) found rates ranging from 43 percent to 69 percent for these indicators in national studies using Medicare claims data.

- The threshold level will be set at 90 percent for physician visits every six months for beneficiaries with chronic stable angina, COPD, CHF, or diabetes.

This threshold is higher since the study by Asch *et al.* (2000) found that six-month visit rates were already above 90 percent for each of these chronic diseases. The threshold is not set even higher since this indicator is primarily intended to detect undertreatment if it occurs.

- The threshold level for meeting the quality target for the ACSC indicator will be set at the FFS national average rate per 1,000 Medicare beneficiaries.

For example, if the national average rate for a performance year is found to be 33.3 ACSC admissions per 1,000 Medicare beneficiaries, as was found in an HER study for the Medicare+Choice population (McCall *et al.*, 2001), then any rate below that figure would mean the PGP had met this target. (In contrast to the other quality indicators, where higher results are better, for ACSCs lower rates represent better performance.)

There are two reasons for using the national average figure as the threshold for the ACSC quality indicator. First, excessively low rates of ACSCs may indicate access problems and not quality improvement. Thus while a 100 percent performance target may be desirable for other quality indicators, such as eye exams for diabetics, a less extreme approach is warranted for the ACSC indicator. Second, even this modest, national average threshold will save money for Medicare while at the same time

encouraging improvements in quality. Reductions in ACSC admissions will directly and immediately benefit CMS due to the relatively high costs associated with hospital admissions. At the same time, those forgone costs will also aid PGPs in meeting their cost targets under this demonstration.

- A lower-bound patient satisfaction threshold will be determined which will serve to detect undertreatment if it exists.

The threshold level for the patient satisfaction indicator is still under consideration. It will be set based on the survey instrument selected and available data on patient satisfaction outcomes for PGPs and Medicare beneficiaries using that measure.

We also considered setting minimum target levels for all of the indicators which would be required for PGPs to achieve in order to earn a bonus. This would enable the demonstration to avoid paying bonuses to PGPs that can be considered low quality providers, at least as measured by these indicators. However, we believe that minimum target levels should not be required for this demonstration for two reasons. First, substantial numbers of beneficiaries will experience health care improvements if low performing groups increase quality in response to the financial incentives provided by this demonstration. Second, we recommend that larger improvements be required to achieve the quality improvement targets for PGPs starting at lower levels of quality.

Substantial Improvement-Over-Time Targets

- For nine of the eleven recommended indicators (all except the ACSC and patient satisfaction indicators), quality improvement over time will be demonstrated if the PGP's performance improves by 10 percent of its quality "deficit" from the previous year. The quality deficit is defined as the ideal rate (100%) minus the actual rate.

For example, if 50 percent of a PGP's diabetics had hemoglobin A1c tests in one year, it would have to raise that level to 55 percent the following year to demonstrate it had met the quality improvement target for that indicator. (A 50 percent deficit means 5 percentage points improvement required.) For these nine recommended indicators, the ideal rate should be 100 percent since they are recommended for all Medicare beneficiaries with the indicated condition (absent identified exclusions).

An advantage of this approach, which is also used by Medicare's Quality Assessment and Performance Improvement (QAPI) program is that it requires smaller improvements from PGPs that are already performing relatively well on an indicator and thus will have a harder time improving significantly. (For example, a PGP at 70 percent performance on hemoglobin A1c tests for diabetics would only need to improve to 73 percent in the following year.)

One limitation on this approach is that it is difficult to apply for the three indicators which are measured over two-year periods (eye exams for diabetics, lipid profiles for diabetics, and mammograms). For these indicators, four years of data are required since a two-year performance period must be compared with a previous two-year base period to demonstrate improvement over time. As a result, to limit the data analysis required for the PGP demonstration, we recommend that the substantial improvement over time target only be applied in the third year of the demonstration for the quality indicators measured over two-year periods. In the third year the required four years of claims data will be available (three performance years plus one base year). The

fixed threshold targets will still apply in all three years of the demonstration for those indicators.

For the ACSC indicator an ideal rate has not yet been determined. As a result:

- A 10 percent reduction compared with the previous year's ACSC admission rate will qualify for meeting the quality improvement target.

For example, if a PGP had a rate of 40 ACSC admissions per 1,000 Medicare beneficiaries in the base year, then it will be required to reduce that rate to 36 admissions per 1,000 in the performance year to meet this target. We will evaluate the likelihood of a 10 percent reduction occurring by chance in a PGP demonstration simulation we will conduct later this year, and recommend a larger percentage reduction target if this appears likely to happen. ACSC rates will also be age and gender adjusted to ensure valid comparisons between demonstration years.

Table ES-3 summarizes the threshold and improvement targets for each of the recommended quality indicators.

Recommendation 19: The Portion of the PGP Bonus Pool That Can be Earned Through Quality Performance

- HER recommends that 30 percent of the PGP Bonus Pool be reserved as the portion that can be earned through quality performance.

Table ES - 3

Recommended Quality Performance Targets For Each Quality Indicator

Quality Indicator	Substantial Improvement Over Time Target	Pre-determined Threshold Target
Eye exams every 2 years for diabetics	10 percent improvement over the deficit from 100 percent compliance*	75 percent compliance
Hemoglobin A1c test every year for diabetics	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Lipid profile test every 2 years for diabetics	10 percent improvement over the deficit from 100 percent compliance*	75 percent compliance
Mammogram every 2 years for women aged 52-69	10 percent improvement over the deficit from 100 percent compliance*	75 percent compliance
Chest radiograph and electrocardiogram ≤ 3 months after initial diagnosis of CHF	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Left ventricular ejection fraction test during the current year for beneficiaries hospitalized with a principal diagnosis of CHF during the current year.	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Physician visit every 6 months for beneficiaries with chronic stable angina, COPD, CHF, or diabetes	10 percent improvement over the deficit from 100 percent compliance	90 percent compliance
Annual influenza vaccinations for all beneficiaries age 65 or older	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Pneumonia vaccination status for all beneficiaries age 65 or older	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Rate of ACSC admissions per 1000 Medicare beneficiaries	10 percent reduction from the previous year's rate	National average rate for FFS beneficiaries
Patient satisfaction (specific indicator to be determined)	N/A	To be determined

*For the quality indicators measured over a 2 year period, the substantial improvement over time target will only be available during the third year of the demonstration. At that point, four years of data will be available (3 performance years plus 1 base year) to enable improvement calculations to be performed.

There are three reasons for recommending that 30 percent of the PGP Bonus Pool be reserved for quality improvements. First, there is a mandate to include incentives for quality improvement in the BIPA 2000 legislation, so a significant percentage is warranted. Moreover, the IOM recently called for pilot testing public sector payment mechanisms with explicit quality improvement incentives for health care providers (IOM, 2001a). It noted that very little attention has been paid to development of ways to align payment incentives with quality improvement. The PGP demonstration represents an opportunity to demonstrate how cost saving and quality improvement incentives can be jointly applied in public sector payment mechanisms. Over time, the percentage allocated to quality performance could rise as PGPs become more familiar with these new types of incentives.

Second, the recommended level of 30 percent should provide a sufficient incentive to motivate PGPs to take action to improve quality. If 30 percent of the PGP Bonus Pool is reserved for quality performance incentives, that means that up to 24 percent (30% quality share x 80% Medicare Savings Sharing Rate) of Medicare Savings can be earned through quality performance. HER simulations have shown that this provides for significant bonus payments to be earned by PGPs. At the same time, this approach reserves the majority (70%) of the PGP Bonus Pool for the cost saving incentives that are the primary focus of the demonstration.

Third, Medicare is protected financially since the actual payments made to PGPs based on quality performance will depend on the total amount of Medicare Savings they have achieved. Quality improvement bonuses will not be paid by CMS if Medicare Savings are not generated by PGPs to fund them.

One possible disadvantage of providing significant direct financial incentives for quality improvement is that they are relatively unexplored territory for public sector payment systems. As a result, most PGPs will be unfamiliar with them in this context. This concern is mitigated by the overall design of the demonstration, however, since PGPs can only benefit from its incentives. They will not be subject to financial penalties if their performance does not reach their quality improvement targets.

Recommendation 20: Calculating and Allocating Process and Outcome Quality Improvement Bonuses

- Participating PGPs will earn 1/8th of the quality portion of the PGP Bonus Pool for achieving a target for each of the eight final quality indicators selected for the demonstration.

Our recommended approach enables a PGP to earn its quality improvement bonus in the most flexible way, in eight discrete segments of 1/8th each. Hence, achieving a quality target is accorded equal weight for each indicator. If targets are met for all eight indicators, the PGP will earn the full 30 percent of the PGP Bonus Pool available for quality incentives. If targets are only met for some of the eight indicators, then the percentage of the PGP Bonus Pool earned by the PGP will be reduced proportionately. However, some portion of the PGP Bonus Pool will be earned for each target achieved, so PGPs will have continuing incentives for quality improvement, even if they believe they are not able to meet the targets for some of the indicators. A 1/8th portion represents 12.5% of the quality portion of the PGP Bonus Pool, 3.75% of the total PGP Bonus Pool (12.5% share of quality portion x 30% quality share of PGP bonus pool), and

3% of Medicare Savings (3.75% share of PGP Bonus Pool x 80% Medicare Savings Sharing Rate).

Although a higher weight might be assigned for some of the quality targets (e.g., patient satisfaction), we recommend equal weights for two reasons. First, equal weighting is simple and avoids further complexity for implementing the demonstration. Second, it is not obvious how to differentially weight the quality indicators. Weights could be based on equating the likelihood of achieving a pre-determined rate of quality improvement for each indicator. Those probabilities could be derived from the standard deviations of the indicators. Alternatively, the underlying utility gains that patients assign to pre-determined quality improvements could be assessed, such as the value of raising patient satisfaction 5 percent versus the value of a 5 percent reduction in the risk of blindness through providing eye exams to diabetics. We believe that both approaches are promising and that work should be actively pursued to develop methodologies for assigning weights to different quality indicators. However, in the absence of widely accepted weighting methodologies at the present time, we recommend using equal weights for this demonstration.

Part I

Introduction

I: Introduction

The physician group practice (PGP) demonstration is a unique reimbursement mechanism through which providers are rewarded for coordinating and managing the overall health care needs of a non-enrolled, fee-for-service (FFS) patient population. It offers an opportunity to test whether a different financial incentive structure can improve service delivery and quality for Medicare patients and ultimately prove cost-effective.

The PGP demonstration superimposes new incentives on traditional FFS reimbursement that are more in line with capitation incentives. PGP organizations will have an incentive to reduce utilization for Medicare FFS patients. However, PGPs that do not reduce utilization are not penalized under the demonstration. The PGP demonstration includes explicit incentives for quality improvement. Performance on both process and outcome quality indicators will be used in the calculation of performance bonuses, together with cost savings. The PGP demonstration will also test and refine a range of technical methods for this new reimbursement system. These include procedures for setting performance targets and for measuring Medicare cost savings and quality improvement.

A legislative mandate for the PGP demonstration was included in the Medicare, Medicaid, and State Child Health Insurance Program Benefits Improvement and Protection Act (BIPA) of 2000. We have summarized below the key provisions of that legislation, which relate to this demonstration. In line with the general purpose of this Design Report, the summary focuses on the demonstration design issues in the legislation; administrative mandates are not addressed in detail. The general intent of the PGP demonstration can be summarized from BIPA 2000 as including five goals:

1. To test the use of incentives for health care groups, including physicians and other providers.
2. To encourage coordination of health care furnished under Medicare parts A and B.
3. To encourage investment in administrative structures and processes for efficient service delivery.
4. To reward physicians for improving health care processes and outcomes.
5. To focus on analysis of the efficiencies and advantages of furnishing health care in a group-practice setting as compared to other health care delivery systems.

To achieve these ends the legislation gives the Centers for Medicare & Medicaid Services (CMS) broad discretion on many demonstration design issues. However, BIPA 2000 does mandate a number of design characteristics for the PGP demonstration. They can be summarized as follows:

- 1) The focus must be on incentives to “health care groups,” which are physician groups at a minimum. Hospitals and other providers can be included in the “group”, but this is not required.
- 2) Participating health care groups must agree to be paid on a fee-for-service basis.
- 3) Where deemed appropriate by the Secretary, payment for all services provided by members of the health care group must be made to a “single entity.”
- 4) Beneficiaries of participating health care groups must be “notified of the incentives, and of any waivers of coverage or payment rules applicable to such group under such demonstration.”
- 5) Performance targets applied to health care groups must include “a base expenditure amount, equal to the average total payments under parts A and B for patients served by the health care group on a fee-for-service basis in a base period determined by the Secretary.”

- 6) The performance target for a health care group must reflect its “base expenditure amount adjusted for risk and expected growth rates”.
- 7) A bonus shall be paid to each participating health care group for each year equal to a portion of the Medicare savings realized relative to the performance target.
- 8) An additional bonus for process improvements and patient outcome improvements shall also be paid to each participating health care group for each year equal to a portion of the Medicare savings.
- 9) Bonus payments are limited as necessary to ensure that the aggregate expenditures, inclusive of bonus payments, with respect to beneficiaries within the scope of the demonstration do not exceed the amount the Secretary estimates would be expended if the demonstration were not implemented.
- 10) Participating health care groups receiving payment under the demonstration must agree to accept such payment as payment in full. However, they can collect deductibles and coinsurance from beneficiaries.

In addition to their standard Medicare FFS reimbursement, which they will continue to receive, PGPs participating in the demonstration will be eligible to earn annual performance bonus payments. PGPs can earn bonus payments for both cost savings and quality performance. To earn bonus payments, a PGP must generate positive Medicare savings.⁵ If a PGP does not generate positive Medicare savings, no cost or quality bonuses are paid, and no savings accrue to the Medicare program. A summary of the steps involved in calculating Medicare savings is as follows:

- 1) For beneficiaries assigned to a PGP in the base year, the base year per capita expenditures are calculated.⁶
- 2) The expenditure target for the performance year is calculated as follows:

⁵ To generate positive Medicare savings, the PGP must satisfy a growth rate threshold. See Recommendation 7: Thresholds for Bonus Payments in Part III below.

⁶ For each beneficiary assigned to the PGP in the base year, base year expenditures are annualized by dividing expenditures by the fraction of the year they were enrolled in both Part A and Part B. Base year per capita expenditures are then weighted by this fraction.

Target = Adjusted Base Year Per Capita Expenditures × (1 + Expected Growth Rate)

Per capita expenditures in the base year are adjusted to account for differences in the casemix of beneficiaries assigned to the PGP in the performance year. The expected growth rate is defined as the growth rate in per capita expenditures in the PGP's local market area (comparison group) between the base and performance years, adjusted for casemix change.

- 3) Medicare savings are computed as the difference between the expenditure target and the PGP's per capita expenditures in the performance year (for beneficiaries assigned to the PGP in the performance year),⁷ multiplied by the number of full-time equivalent (FTE) beneficiaries (person years) assigned to the PGP in the performance year.⁸

Medicare Savings = (Target - Performance Year Per Capita Expenditures) × FTE Beneficiaries.

The distribution of Medicare savings in the PGP demonstration is shown in Figure I-1.

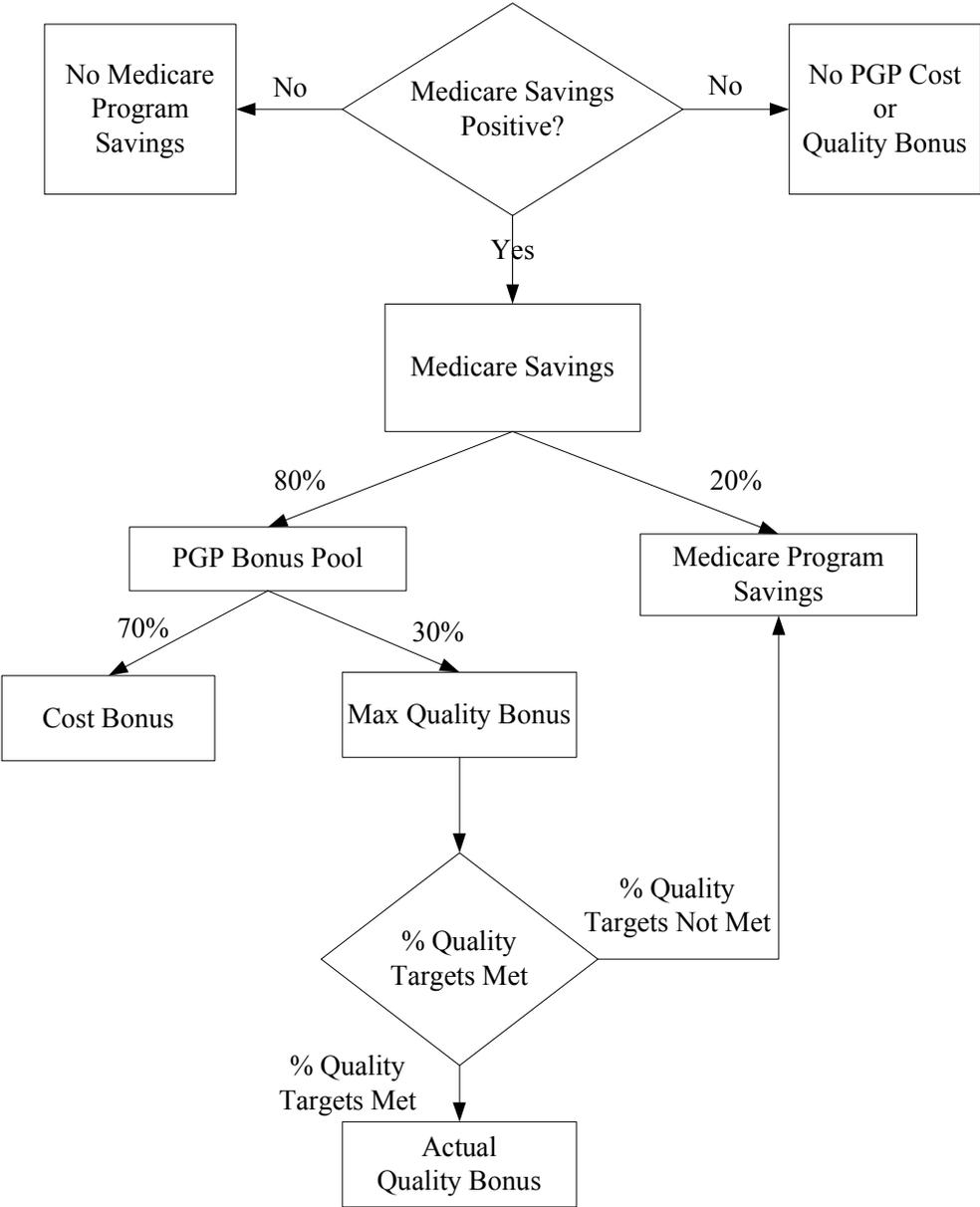
As shown in Figure I-1, if a PGP does not generate positive Medicare savings, no cost or quality bonuses are paid, and no savings accrue to the Medicare program. For PGPs that do generate positive Medicare savings, the recommended sharing rate equals 80%, which means that 80% of Medicare savings is allocated to the PGP bonus pool, and the remaining 20% is savings for the Medicare program. The PGP automatically receives 70% of the PGP bonus pool as a "cost bonus," and can earn the remaining portion of the PGP bonus pool by providing high quality care. The maximum quality bonus the PGP

⁷ For each beneficiary assigned to the PGP in the performance year, performance year expenditures are annualized by dividing expenditures by the fraction of the year they were enrolled in both Part A and Part B. Performance year per capita expenditures are then weighted by this fraction.

⁸ FTE beneficiaries are determined as follows. For each beneficiary assigned to the PGP in the performance year, the fraction of the year the beneficiary was enrolled in both Part A and Part B is calculated. FTE beneficiaries equal the sum of these fractions (i.e., equals number of person years).

Figure I-1

Recommended Distribution of Medicare Savings in the PGP Demonstration



SOURCE: Health Economics Research, Inc.

can earn is the remaining 30% of the PGP bonus pool. The “actual” PGP quality bonus equals the maximum quality bonus multiplied by the percentage of quality targets met by the PGP. The remainder of the maximum quality bonus, if any, is additional savings for the Medicare program.

For PGPs that generate positive Medicare savings, Table I-1 shows the maximum and minimum payments to PGPs and the Medicare program as a percentage of Medicare savings. PGPs will receive a minimum 56% of Medicare savings through their cost bonus (80% sharing rate x 70% cost bonus percentage = 56%). The maximum quality bonus is 24% of Medicare savings if all quality targets are met (80% sharing rate x 30% maximum quality bonus percentage x 100% quality targets met percentage = 24%). The minimum quality bonus is 0% of Medicare savings if no quality targets are met (80% sharing rate x 30% maximum quality bonus percentage x 0% quality targets met percentage = 0%). Thus, PGPs will receive 80% of the Medicare savings when all quality targets are met, versus 56% when no quality targets are met. Conversely, the Medicare program receives 44% of Medicare savings when no quality targets are met, versus only 20% when all quality targets are met.

For PGPs that generate positive Medicare savings, a summary of the steps involved in calculating the PGP bonus, and savings for the Medicare program, is as follows:

- 1) The sharing rate equals 80%, which means 80% of Medicare savings is allocated to the PGP bonus pool, and the remaining 20% is the minimum savings for the Medicare program:

Table I-1

**PGP Bonus and Medicare Program Savings
as a Percentage of Demonstration Medicare Savings**

	<u>All Quality Targets Met</u>	<u>No Quality Targets Met</u>
PGP Bonus	80%	56%
Cost	56	56
Quality	24	0
Medicare Program Savings	20	44

NOTE: Assumes positive Medicare savings. Otherwise no cost or quality bonuses are paid, and no savings accrues to the Medicare Program.

SOURCE: Health Economics Research, Inc.

$$\text{Bonus Pool} = \text{Medicare Savings} \times 80\%$$

$$\text{Minimum Medicare Program Savings} = \text{Medicare Savings} \times 20\%$$

- 2) The cost bonus equals 70% of the bonus pool:

$$\text{Cost Bonus} = \text{Bonus Pool} \times 70\%$$

- 3) The remaining 30% of the bonus pool is the maximum quality bonus the PGP can earn:

$$\text{Maximum Quality Bonus} = \text{Bonus Pool} \times 30\%$$

- 4) The “actual” quality bonus earned by the PGP equals the maximum quality bonus multiplied by the percentage of quality targets met by the PGP. The remainder of the maximum quality bonus is additional savings for the Medicare program:

$$\text{Quality Bonus} = \text{Maximum Quality Bonus} \times \text{Percentage of Quality Targets Met}$$

$$\text{Additional Medicare Program Savings} = \text{Maximum Quality Bonus} \times (100\% - \text{Percentage of Quality Targets Met})$$

- 5) The total PGP bonus (cost plus quality), and savings for the Medicare program, are as follows:

$$\text{Total Bonus} = \text{Cost Bonus} + \text{Quality Bonus}$$

Medicare Program Savings = Minimum Medicare Program Savings +
Additional Medicare Program Savings

To prepare for implementation of the PGP demonstration, as provided for under BIPA 2000, this Design Report includes recommendations regarding 20 key design issues. Rationales and discussion are included for each recommendation. The recommendations are organized into four broad categories: Demonstration Eligibility and Assignment; Bonus Computation; Expenditure Definition and Adjustments; and Quality Targets and Bonuses.

Part II

Demonstration Eligibility and Beneficiary Assignment

II: Demonstration Eligibility and Beneficiary Assignment

Recommendation 1: Eligibility of PGPs for the Demonstration

The PGP demonstration should include several different types of PGPs in order to test its new incentives in a range of organizational and clinical environments. In particular, HER recommends that the demonstration be open to:

- Stand-alone PGPs as well as those affiliated with hospitals, other provider organizations, or integrated delivery systems (IDSs); and
- PGPs with national reputations as tertiary referral centers as well as smaller, regional providers.

However, HER also recommends that four more restrictive eligibility criteria also be applied. They relate to the capability of participating PGPs to respond effectively to the demonstration's new incentives. In particular, the demonstration should be restricted to:

- Large PGPs, with at least 200 physicians and at least 15,000 to 20,000 Medicare FFS beneficiaries assigned;
- Multispecialty PGPs;
- PGPs with relatively well developed information, clinical, and management systems; and
- Contracting entities which are either stand-alone PGPs or distinct organizational units defined as physician groups within larger IDSs or other multi-provider organizations.

Discussion: Large PGPs are necessary in order to ensure that the demonstration participants will have the administrative and clinical capabilities necessary to respond to the demonstration's new incentives. Large PGPs will also better ensure that sufficient numbers of Medicare beneficiaries are treated by the PGPs in order to provide for statistical stability in calculation of budget targets and performance comparisons. Based

on our simulations and analysis (see Recommendation 7 below), we estimate that 20,000 beneficiaries will be the minimum number needed for some PGPs. However, the minimum number of beneficiaries required per PGP will vary depending on the number of Medicare beneficiaries in the PGP's market comparison area. A large market comparison area will provide for statistical stability in budget performance comparisons (which are based on market area growth rates) for PGPs with somewhat fewer Medicare beneficiaries. However, HER recommends that 15,000 Medicare beneficiaries be the overall minimum number expected for participating PGPs, regardless of the size of the market comparison area. Given the requirement of 200 physicians for participating PGPs, the expectation of at least 15,000 to 20,000 Medicare beneficiaries assigned to each PGP should generally be achieved. For example, assuming that each physician's panel includes 125 Medicare beneficiaries, a PGP with a minimum of 200 physicians would treat 25,000 patients. However, that 25,000 figure may not be achieved since the recommended beneficiary assignment criteria for this demonstration provide for a beneficiary to be assigned to a PGP only when the PGP provides at least one evaluation and management (E&M) service to the beneficiary, and more E&M services than any other physician practice (see Recommendation 2 below). Nonetheless, we expect the requirement of 200 physicians will provide a sufficient number of assigned beneficiaries.

A demonstration simulation study to be conducted over the next several months will enable us to estimate the average beneficiaries to PGP physician ratio more precisely when applying our recommended beneficiary assignment criteria. In addition, most PGPs will not be aware of the number of unique Medicare beneficiaries they serve, especially under our recommended assignment criteria. Analysis of Medicare claims

data would be needed to evaluate the eligibility of most PGPs with regard to the number of assigned beneficiaries they serve, and this would not be feasible during the limited time available to CMS between receiving applications for the demonstration from PGPs and the point at which selection of participating PGPs would need to be made.

As a result, HER recommends that the PGP size eligibility criterion focus on the requirement of at least 200 physicians. That figure is much easier to measure and can be expected to provide a sufficient number of beneficiaries for participating PGPs. Smaller groups might be able to implement the demonstration, but eligibility for them should be deferred until after the demonstration has been concluded and the larger groups have shown that the new incentives can be utilized by physician groups to improve health care efficiency and quality.

The most recent data available from the Medical Group Management Association (MGMA) and the American Medical Group Association (AMGA) indicate that there should be sufficient numbers of eligible PGPs for the demonstration with a requirement of at least 200 physicians per PGP. Table 1-1 below includes the membership figures for the MGMA for 1999. Table 1-2 below includes the membership figures for the AMGA for 2001.

Table 1-1

MGMA Membership by Group Size, 1999

<u>Group Size (Number of Physicians)</u>	<u>Number of MGMA Member Groups</u>
51-75	143
76-150	151
151+	203

SOURCE: Medical Group Management Association

Table 1-2

AMGA Membership by Group Size, 2001

<u>Group Size (Number of Physicians)</u>	<u>Number of AMGA Member Groups</u>
100-150	32
150-200	11
200+	44

SOURCE: American Medical Group Association

There is likely some overlap in the PGPs listed in Tables 1-1 and 1-2, since membership in the AMGA and MGMA is not mutually exclusive. Moreover, MGMA does not routinely report data on groups of 200 or more. Nonetheless, these data indicate that there are at least 75 to 100 PGPs with at least 200 physicians that are members of one or both of these organizations, which should be sufficient for purposes of this demonstration.

The requirement for participating PGPs to be multispecialty groups is consistent with the expectation that PGPs participating in the demonstration must possess the capacity to respond to the demonstration's incentives by better coordinating care across multiple providers and multiple sites of care. This requirement could be defined by the total number of specialties provided by the PGP, or by setting separate criteria for primary care, medical specialties, and surgical specialties. Most PGPs meeting the criterion for the minimum size of 200 physicians will include a sufficient range of different specialties, but a formal set of criteria should be established to avoid any ambiguity in the solicitation. We recommend the following definition be applied to define multispecialty groups eligible for the PGP demonstration: "At least one primary care specialty relevant to Medicare beneficiaries (internal medicine, general practice, preventive medicine, geriatric medicine, or family practice) and at least five other medical or surgical specialties, but with at least 15 percent of the total number of physicians serving in a primary care specialty."

The requirement for participating PGPs to have relatively well developed information, clinical, and management systems is also consistent with the expectation that PGPs participating in the demonstration will possess the capacity to better coordinate

care and to respond effectively to the demonstration incentives. HER recommends applicant PGPs should be required to present a plan describing how their systems will be used to respond to the demonstration incentives. PGPs with less developed systems should be largely screened out through including this requirement in the solicitation

However, PGPs with systems under development should be eligible to participate, if they have clearly articulated plans for applying their new systems to respond to the demonstration incentives. This type of PGP should, however, be required to demonstrate the operational credibility of its plans. This criterion may be difficult to evaluate quantitatively, but applicants should be required to describe their systems development efforts in detail and to describe how they plan to improve coordination of care and improve their cost and quality outcomes. Site visits may be needed to some PGPs to fully evaluate developing systems and the credibility of their plans.

In theory, a range of different levels of internal PGP systems development might enable better testing of the importance of this factor for the ability of groups to respond to the demonstration incentives. However, since the PGP demonstration is breaking new ground, with new types of incentives for FFS practice, it is prudent to focus this initial effort on PGPs which are more likely to be able to respond effectively.

The requirement that contracting entities for the demonstration be either stand-alone PGPs or distinct organizational units defined as physician groups within IDSs or other multi-provider organizations is based on the intent of BIPA 2000 that the demonstration investigate the advantages of providing health care in a group practice setting. Moreover, BIPA 2000 indicates that the PGP demonstration should develop

contracts for payment to a “single entity,” although this is qualified as “where determined appropriate by the Secretary.”

However, the range of hybrid organizational entities, affiliations, and ownership structures among PGPs and other health care facilities and providers has expanded greatly in recent years. So it is possible that many different types of organizations may respond to the demonstration’s solicitation. These could include IPAs, PHOs, MSOs, PSOs, ISDNs, and others. In addition, these new organizational forms (and the terminology used to describe them) have not yet been standardized and remain in considerable flux (Kongstvedt, 2001). Some iterative discussion with applicants may be necessary during the review process to ensure that the demonstration incentives will directly affect the physicians involved.

Interdependencies: The threshold for bonus payment (Recommendation 7) depends for statistical stability on the number of Medicare beneficiaries served by a PGP and the number of Medicare beneficiaries in its market comparison area. Our recommendation of a 2 percent savings threshold for actual performance versus a PGP’s target is the basis for the recommended minimum of 15,000 to 20,000 beneficiaries discussed above, to better ensure that the “savings” generated by a PGP under this demonstration represent true savings for Medicare, and not the result of random statistical fluctuations. If the bonus threshold is changed to a figure other than 2 percent, then the minimum number of beneficiaries needed per participating PGP may also need to change, depending on the size of the PGP’s market comparison area.

As noted, the beneficiary assignment criteria (Recommendation 2) will affect the calculation of the number of beneficiaries served by each PGP for purposes of this demonstration. HER's recommendation is that beneficiaries be assigned to a PGP if they receive at least one E&M service from the PGP, and more E&M services from that PGP as from any other physician practice. If these assignment criteria are changed, they may affect the ability of some PGPs to reach the recommended minimum of 15,000 to 20,000 assigned Medicare beneficiaries.

Recommendation 2: Assigning Beneficiaries to PGPs

- Beneficiaries who receive at least one Evaluation and Management (E&M) service from a participating PGP are eligible for assignment to the PGP. Beneficiaries with any Medicare+Choice enrollment are not eligible for assignment. Beneficiaries who receive more E&M services (measured by Medicare payments) from a participating PGP than from any other physician practice (group or solo) will be assigned to the PGP. Certain E&M services (e.g., emergency department visits) will be excluded from assignment. Beneficiaries will be assigned to at most one participating PGP.

Background: Since the PGP demonstration is a fee-for-service innovation, there is no enrollment process whereby beneficiaries accept or reject participation. Thus beneficiaries need to be “assigned” to PGPs based on utilization of Medicare-covered services. The intention of the demonstration is to give each participating PGP an incentive to manage the health care of the beneficiaries assigned to it.

Discussion: The ability of a physician practice to manage the health care of a beneficiary depends on its control over the beneficiary's utilization of services. Because of this, a participating PGP providing the largest share of E&M services is, we believe, in

the best position to manage the health care of the beneficiary. The data processing steps involved in assigning beneficiaries to PGPs are discussed in Appendix A. Unique assignment to PGPs prevents CMS from paying bonuses more than once when multiple PGPs serve overlapping Medicare patient populations. Also, because assignment is based only on E&M procedure codes, assignment does not depend on physician specialty.

Certain E&M services such as consultations and emergency department visits do not reflect the PGP's ability to manage the health care of the beneficiaries assigned to it, and thus should not be used for assignment. E&M services included and excluded in the assignment criteria are shown in Table 2-1.⁹

Certain beneficiaries are problematic for any assignment criteria. For example, snowbirds are Medicare beneficiaries who spend winters in one area of the country and summers in another. PGPs that are assigned snowbirds might not feel primary responsibility for their health care. Another example is long-distance referrals. Other things equal, a PGP is likely to have less control over the health care of certain beneficiaries like snowbirds and long-distance referrals. It is not however necessary to exclude these beneficiaries from the demonstration, the reason being that the beneficiary assignment criteria requires that a PGP provide more E&M services than any other physician practice. As discussed below, we expect that a PGP will feel ownership of the majority of patients assigned to it under the demonstration. Further, exclusion of these

⁹ For a complete listing of E&M procedure codes, see Current Procedural Terminology 2001, American Medical Association, AMA Press, 2000.

Table 2-1

**Evaluation and Management Services
Included and Excluded in Beneficiary Assignment Criteria**

Included

Office or Other Outpatient Services
Hospital Observation Services
Hospital Inpatient Services
Nursing Facility Services
Domiciliary, Rest Home (e.g., Boarding Home), or Custodial Care Services
Home Services
Prolonged Services
Case Management Services
Care Plan Oversight Services
Preventive Medicine Services

Excluded

Consultations
Emergency Department Services
Patient Transport
Critical Care Services
Neonatal Intensive Care
Newborn Care
Special Evaluation and Management Services
Other Evaluation and Management Services

NOTE: For a complete listing of Evaluation and Management Services,
see Current Procedural Terminology cpt 2002, American Medical Association, AMA Press, 2001

SOURCE: Health Economics Research, Inc.

beneficiaries might not be practical. For example, CMS data do not permit ready identification of snowbirds. Specifically, the time period(s) that beneficiaries spend outside of their permanent residence area cannot be ascertained through CMS's Enrollment Data Base (EDB). This is because many, if not the vast majority, of Medicare beneficiaries do not file changes of address with CMS when they vacation elsewhere for extended periods of time.

In an HER project for CMS entitled *Research and Analytic Support for Implementing Performance Measurement in Medicare Fee for Service*,¹⁰ 1997 data on four PGPs is used to perform simulations on criteria for assigning Medicare fee-for-service beneficiaries to these PGPs. Table 2-2 presents results from that project. Table 2-2 shows the number of beneficiaries assigned to each of the four PGPs under various assignment criteria. Column 1 shows the number of beneficiaries assigned under the criteria that the PGP provides any physician services to the beneficiary. Column 2 shows the number of beneficiaries assigned under the criteria that the PGP provides at least one primary care E&M service to the beneficiary.¹¹ Finally, Column 3 shows the number of beneficiaries assigned under the criteria that the PGP provide at least as many primary care E&M visits as any other physician practice. Note that the assignment criteria in Column 3 is similar to that recommended for the PGP demonstration, the main difference

¹⁰ See McCall, Pope, Adamache, *et al.*, *Research and Analytic Support for Implementing Performance Measurement in Medicare Fee for Service*, CMS Contract No. 500-95-0058, First Annual Report, November 17, 1998.

Also see McCall, Pope, *et al.*, *Research and Analytic Support for Implementing Performance Measurement in Medicare Fee for Service*, CMS Contract No. 500-95-0058, Second Annual Report, January 6, 2000.

¹¹ Primary care E&M services as defined in McCall *et al.*, 1998 are: office or other outpatient services; nursing facility services; domiciliary, rest home, or custodial care services; home services; care plan oversight services; and end stage renal disease services.

Table 2-2

Number of Beneficiaries Assigned to Physician Group Practice (PGP), 1997

<u>PGP</u>	(1) Beneficiaries With a Claim at the <u>PGP</u>	(2) Beneficiaries With Any Primary Care at the <u>PGP</u>	(3) Beneficiaries That Received Equal or Most of Their Primary Care Care at the <u>PGP</u>	Percent Retained in Going From:		
				<u>(1) to (2)</u>	<u>(2) to (3)</u>	<u>(1) to (3)</u>
Group A	76,570	55,854	32,945	72.9%	59.0%	43.0%
Group B	58,765	49,292	37,417	83.9	75.9	63.7
Group C	35,257	29,034	20,264	82.3	69.8	57.5
Group D	9,776	8,516	4,408	87.1	51.8	45.1
Total	180,368	142,696	95,034	79.1	66.6	52.7

(1) Equal or highest proportion of total annual relative value units provided by the group practice.

SOURCE: NT McCall, GC Pope, KW Adamache, *et al.* "Research and Analytic Support For Implementing Performance Measurement In Medicare Fee For Service",

being that assignment for the PGP demonstration is based not only on primary care E&M services, but on non-primary E&M services as well. Therefore, results in Table 2-2 are relevant for the PGP demonstration assignment criteria.

In Table 2-2, the percentage of beneficiaries retained in going from the first assignment criteria to the second is 79.1%, and the percentage of beneficiaries retained in going from the first assignment criteria to the third is 52.7%. Thus, the PGP demonstration assignment criteria will result in only about half as many beneficiaries assigned to participating PGPs as compared to the assignment criteria that the PGP provide any physician services to the beneficiary. Nevertheless, except for the one relatively small PGP, the number of beneficiaries assigned to each PGP ranges from about 20,000 to 40,000. This exceeds HER's minimum recommended sample size of 15,000 (Recommendations 1 and 7). Note that under the PGP demonstration assignment criteria the number of assigned beneficiaries for each PGP would be greater than Column 3 because the assignment criteria is based not only on primary care E&M services, but non-primary E&M services as well.

To examine the degree to which physicians within a participating PGP will feel ownership of the patients assigned to them under the PGP demonstration, we again use results from McCall *et al.*, 1998. As discussed above, the assignment criteria used in that study is similar to the assignment criteria recommended for the PGP demonstration. HER provided each PGP with a sample of 120 patients that had been assigned to them. We then asked the physician within the PGP who had provided the most care to the beneficiary whether they considered the patient to be their patient. We analyzed the degree to which physicians felt ownership of patients assigned to them by giving them

three potential responses: definitely, maybe, or not their patient. Almost three-quarters of physicians within the four PGPs felt definite ownership of the patients assigned to the PGP. Another approximately 10% felt that the patients maybe were their patient, and almost 20% did not feel any ownership of the patients assigned to them. These results suggest that PGP demonstration participants will feel ownership of most of the beneficiaries assigned to them.

We are still considering options for several scenarios. First, a beneficiary may receive an equal amount of E&M services from two participating PGPs during a year, and no E&M services from any other physician practice (group or solo). Under this scenario, the beneficiary would not be assigned to either PGP under the recommended assignment criteria. Second, and potentially more important, is the possibility that a PGP provides the majority of Medicare-covered services to a beneficiary during a year, and yet the beneficiary is not assigned to the PGP under the recommended assignment criteria. Assignment criteria may be modified after further analysis of these situations.

Part III

Bonus Computation

III. Bonus Computation

Recommendation 3: Comparison Population and Expected Growth Rate

- The comparison population for a participating PGP is fee-for-service Medicare beneficiaries residing in the PGP's market area that are not assigned to the PGP. The PGP's market area will be defined as counties in which 1% or more of the beneficiaries assigned to the PGP reside. The market area will be defined for both base and performance years, and may differ between the two years to reflect changes in the PGP's service area (although in general we expect PGP service areas to be stable over time). The PGP's expected expenditure growth rate will be the change in market area per capita expenditures from the base to the performance year. Market area per capita expenditures will be defined as weighted average county per capita expenditures of market area counties. The weights will be the share of participating PGP beneficiaries residing in each market area county.

Background: The PGP demonstration is designed to reward participating PGPs for improved cost efficiency by comparing cost performance relative to a target. The target will be defined by updating the participating PGP's adjusted¹² base expenditure amount by the rate of cost increase in a comparison population:

Target = (adjusted base year PGP per capita expenditures) * (growth rate in adjusted per capita expenditures in comparison population from base to performance year).

For example, if the adjusted base year PGP per capita expenditures are \$5,000 and the comparison population expenditure growth rate is 5%, then the target is $\$5,000 * 1.05 = \$5,250$. The site and comparison population expenditures both will be adjusted to remove the effects of changes in health status casemix. The Medicare savings and PGP bonuses will be determined by comparing performance year participating site per capita expenditures to the site's target.

¹² Base year expenditures will be adjusted for health status casemix in the performance year.

Discussion: The purpose of the comparison population in the PGP demonstration is to predict what the per capita expenditures of PGP beneficiaries would have been in the absence of the demonstration. That is, what per capita expenditures would be if the cost control performance of the participating PGP is the same as the "market". Then the participating PGP will earn a bonus if and only if its cost control is better than average in its market.

Growth in per capita expenditures is likely to be 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 the local medical care industry. For this reason, local Medicare beneficiaries not assigned to the participating PGP are a natural comparison group for the PGP demonstration. With this comparison group, the participating PGP will earn a bonus if it performs better than its local competitors or market, who face similar market conditions.

A convenient way to define a PGP's market area is by the residence location of beneficiaries assigned to it. This "patient origin" approach to market definition has been widely studied and recommended for hospitals (Zwanziger, Melnick, and Mann, 1990; Zwanziger, Melnick, and Eyre, 1994; Garnick, Luft, Robinson, and Tetreault, 1987), and appears to apply equally well to identifying physician group practice market areas. Defining market areas by patient origin is a sound, empirically-based method that reflects the actual catchment area of each participating PGP. It does this more accurately than pre-specified administrative units such as the county, Metropolitan Statistical Area, or state, or pre-specified geographic radiuses such as 15 miles from the practice location.

The market areas of multi-location practices are accurately identified with the patient origin approach.

In a companion study by HER researchers, Klosterman, Pope, and Kautter (2002) simulated alternative patient-origin-based market areas and target expenditure computations using actual data on beneficiaries assigned to four PGPs. Seven market area definitions were evaluated:

1. counties cumulatively containing at least 70% of a PGP's assigned beneficiaries;
2. counties cumulatively containing at least 80% of a PGP's assigned beneficiaries;
3. counties cumulatively containing at least 90% of a PGP's assigned beneficiaries;
4. counties where at least 5% of a PGP's assigned beneficiaries reside;
5. counties where at least 3% of a PGP's assigned beneficiaries reside;
6. counties where at least 1% of a PGP's assigned beneficiaries reside;
7. counties where at least 5% of a PGP's assigned beneficiaries reside or where the PGP has a Medicare fee-for-service market share of at least 5%.

These market area definitions were evaluated on criteria including:

- defining a valid comparison group for PGP beneficiaries, i.e., it includes the regions where most of the PGP assigned beneficiaries reside.
- sufficient comparison population sample size/stable target growth rate computation
- market area geographically compact/contiguous and has face validity
- minimizes data collection/computation burden (by including fewer counties)

- satisfies first four criteria for a wide range of types of PGPs (e.g., those with snowbirds, urban, rural, large, small, PGPs drawing from a wide range of counties and those drawing from only a few counties).

Based on these criteria, HER concluded that counties where at least 1% of a PGP's assigned beneficiaries reside was the best market area definition, and that is what HER recommends.

The recommended market area definition could define markets including counties in more than one state. This did not happen in our simulations with four actual PGPs, but it could happen if a PGP is located near or on a state boundary, or has practice locations in multiple states. We do not see multi-state market areas as a problem or weakness of the recommended method, but as a strength, because it reflects the locations that a PGP is actually drawing patients from. The recommended method could also possibly define a market area containing non-contiguous counties—for example, a very sparsely populated county lying between two more populous included counties might not be included simply because very few Medicare beneficiaries lived there. Or if a PGP has widely separated practice locations, its market area might not consist of contiguous counties. Again, we do not see this as a weakness of the recommended method, but rather an accurate reflection of the actual catchment area of the PGP. In our simulations with data on four actual PGPs, all market areas defined by the recommended method were contiguous and relatively compact (which was one of the reasons we preferred this method).

County residence of Medicare beneficiaries is available on the Medicare enrollment files. Counties will be included in the market area if they contain 1% or more of beneficiaries assigned to a participating PGP. Counties with less than 1% of

beneficiaries are excluded because they are not part of a PGP's core market area. Beneficiaries in these counties are often "snowbirds" or travel to the practice to receive a specific service. Including all counties with PGP beneficiaries would eliminate the need to define market areas, but would increase administrative burden by requiring data to be processed for large numbers of counties with very few beneficiaries. Also, including distant counties with few beneficiaries in a PGP's comparison group does not have face validity. For example, it is not plausible that Florida counties should be in the market area of a PGP located in Pennsylvania, but patient origin data will show some Florida beneficiaries assigned to a Pennsylvania practice.¹³

Once the market area is defined, the comparison group expenditure growth rate will be computed using weighted averages of county adjusted per capita expenditures. The weight will be the share of PGP beneficiaries in each county comprising the market area. For example, let counties A, B, and C comprise the market area. Suppose 50%, 30%, and 20%, respectively, of total PGP beneficiaries in the market area reside in each county in the base year, and 55%, 35%, and 10%, respectively, in the performance year. Then base year and performance year comparison group adjusted per capita expenditures will be computed as a weighted average of the adjusted per capita expenditures of counties A, B, and C. A weight of 40%, 30%, and 20%, respectively, will be put on each county in the base year, and 55%, 35%, and 10%, respectively, in the performance year. The comparison group (i.e., target) growth rate will then be the growth rate from the base to the performance year of weighted average county adjusted per capita expenditures.

¹³ Presumably these are snowbirds, who reside part of the year in Florida, and part in Pennsylvania. Or they may be beneficiaries who have retired to Florida, but still obtain their medical care in Pennsylvania.

Per capita county expenditures in the market area are weighted by PGP beneficiary frequency for three reasons. The first is that expenditure growth rates may differ among counties; weighting by PGP beneficiary frequencies ensures that the comparison group growth rate will reflect expenditure growth in the counties where most of the PGP's beneficiaries reside. In the example of the preceding paragraph, county A receives a higher weight because it contains more of the PGP's beneficiaries. Second, the distribution of PGP beneficiaries among high- and low-cost counties may change from the base to the performance year. Using year-specific base and performance year weights ensures that changes in the geographic distribution of PGP beneficiaries will be reflected in its performance target. In the example above, suppose the PGP is opening more practice locations in the high-cost urban core county A, while scaling back operations in rural county C. This shift, which will cause the PGP's expected per capita expenditures to grow, is captured by the increased weight put on county A, and the decreased weight on county B, in the performance year. Third, because it is patient-origin-based, the PGP's market area may change from year to year. Some counties may be slightly under the threshold for inclusion in one year, then exceed the threshold in the next year. If a newly-included county is a populous urban county with high costs, for example, it could substantially affect market area average expenditures in the year it is included. Weighting by proportion of PGP beneficiaries in a county will improve the stability of the market area average expenditures. Counties that are borderline for inclusion in the market area will contain very small proportions of PGP beneficiaries, and will receive little weight in the market area average. Hence, the market area average expenditures,

and the target growth rate, will not be very sensitive to whether borderline counties are included in the market area or not.

In HER's companion study (Klosterman, Pope, and Kautter, 2002), we simulated weighting target growth rate computations by share of market area PGP beneficiaries in a county versus by share of market area total number of Medicare fee-for-service beneficiaries in a county. We concluded that the former, which is what we recommend, was preferable. For example, the recommended method produced more stable target growth rate computations across alternative market area definitions.

Several categories of beneficiaries will be excluded from the comparison group. These include:

- beneficiaries with any managed care enrollment during the year;
- beneficiaries assigned to the participating PGP; and
- beneficiaries not using at least one of the evaluation and management services used to assign beneficiaries to a participating PGP.

Beneficiaries enrolled in managed care will be excluded because the fee-for-service claims data necessary to compute expenditure growth are not available for them. The PGP's own beneficiaries are excluded so that the expenditure target is largely independent of the performance of the PGP.¹⁴ Nonusers of the specified evaluation and management services used to assign PGP beneficiaries are excluded to make the comparison population more similar to the population assigned to the PGP. A

¹⁴ Some beneficiaries not assigned to the PGP may nevertheless receive services from it. (Beneficiaries receiving services from the PGP are not assigned to it if they receive a larger share of their evaluation and management services from another provider.) So the comparison group will not be completely independent of the PGP. But any "endogeneity" is likely to be very small.

beneficiary must use at least one of the specified evaluation and management service to be assigned to a participating PGP (see Recommendation 2).

Once these exclusions are made, an effective sample size of at least 15,000-20,000 beneficiaries must remain in the comparison group to ensure the statistical validity

of bonus computations (see Recommendation 7). The effective sample size of the comparison population will differ from its nominal size, because of the weighting of county per capita expenditures by the distribution of PGP beneficiaries (see Appendix F for further discussion and formulas for computing effective sample size). If the PGP frequencies put a high weight on a less populous county within the market area, the effective size will be smaller than the nominal size. If the effective sample size of the comparison group is too small for any reason, either the market area counties will be reweighted¹⁵ or counties will be added to the market area in descending order of number of PGP-assigned beneficiaries until the effective number of comparison beneficiaries exceeds 15,000. CMS and its demonstration technical support contractor will review the final market area and comparison population for reasonableness. If for any reason the comparison population is skewed or unrepresentative, the market area will be modified (e.g., enlarged) so that it defines a representative comparison population.

The comparison population will not consist of the same beneficiaries in both the base and the comparison years. Some beneficiaries will die or move away from the base to the performance years, and others will newly enroll in Medicare or move into the area.

¹⁵ For example, greater weight could be put on a populous urban county in the market area, and less weight on a sparsely populated suburban or rural county, to raise the effective sample size.

In addition, the service area of the participating PGP may expand, contract, or change, which will cause a corresponding change in its patient-origin-based market area. Rather, the comparison population will consist of beneficiaries living in the participating PGP's market area each year. This is similar to the PGP's assigned beneficiary population, which will not consist of the same individuals in the base and performance years either.

The comparison population is designed to be an accurate and independent "peer group" for the beneficiaries assigned to the participating PGP in each year. The comparison population will be monitored by CMS or its contractor for any unusual changes from the base to the comparison year (e.g., a large change in managed care penetration in the market area that dramatically reduces the number of fee-for-service beneficiaries available for the comparison group). Except in unusual cases of major changes in a PGP's practice locations, we expect the PGP's market area and comparison population to be quite stable from year to year. But in the unlikely event that there is a major change in the PGP's organizational structure, the flexible, patient-origin-based approach to market area definition will accurately incorporate changes while maintaining the validity and stability of the bonus computations.

The participating PGP expenditure growth rate will be calculated using all assigned beneficiaries, whether they reside in the PGP's market area or not. Because the market area will include all counties with at least 1% of PGP beneficiaries, the proportion of out-of-area beneficiaries is expected to be small.¹⁶ Further, because beneficiaries will be assigned to the PGP only if they receive at least as many specified evaluation and

¹⁶ However, practices that treat a large proportion of snowbirds, such as those located in Florida or Arizona, may have a considerable proportion of out-of-area beneficiaries.

management services at the PGP as at any other provider, even out-of-area beneficiaries will generally be receiving a significant share of their Medicare services at the PGP. The PGP physicians will control additional services through referrals and recommendations, even if not in-area. For these reasons, it is reasonable to include out-of-area beneficiaries in calculating the PGP expenditure growth rate, so that adequate sample size to ensure an accurate estimate is maintained.

The specific data processing steps involved in defining the market area, and computing the target growth rate, are discussed in Appendix B.

Recommendation 4: Sharing Rate

- The sharing rate is the maximum proportion of the Medicare savings generated by a PGP that is paid to the PGP as a bonus. The sharing rate needs to be high enough to give PGP participants sufficient incentive to participate in the demonstration, but low enough so that the Medicare program shares significantly in any savings. A sharing rate of 80% percent is recommended. With this sharing rate, the PGP earns up to 80% percent of the Medicare savings it generates (depending on its performance with regard to the quality of care targets), and a minimum of 20% of Medicare savings accrues to the government.

Background: The PGP demonstration is intended to give a participating PGP an incentive to generate Medicare savings by lowering the volume of services provided to the beneficiaries assigned to it. Given that the PGP meets all of the quality of care targets established for the demonstration, the bonus paid to the PGP equals the product of the Medicare savings it generates and the sharing rate. For example, if the PGP generates \$8,000,000 in Medicare savings during a performance year of the demonstration, and the sharing rate is 80 percent, then the PGP bonus earned (pre-withhold bonus) is

\$6,400,000. The remaining \$1,600,000 accrues to the government. Both the participating PGP and the government "win" under this example, which is the intention of the PGP demonstration.

Discussion: The sharing rate needs to be set high enough to give PGPs sufficient incentive to participate in the demonstration. To determine a sharing rate that meets this criterion, we simulate bonus payments and revenue changes under various sharing rates. The simulation requires that we make assumptions about PGP demonstration parameters, as well as factors related to the PGP and its market area. Some assumptions are necessarily speculative since they involve future Medicare trends, policy decisions that are not finalized, and typical characteristics and behavioral responses of potential participating PGPs. However, we think our assumptions are reasonable.

The parameter values used in the sharing rate simulation are shown in Table 4-1.

Table 4-1

Parameter Values for Sharing Rate Simulation

<u>CMS Parameters</u>	
Cost Savings Share of Bonus Pool	70%
Quality Performance Share of Bonus Pool	30%
Withhold Rate	25%
Bonus Threshold	2%
<u>PGP Parameters</u>	
Adjusted Base Year Expenditures	\$6,500
Medicare Savings Rate	10%
Assigned Beneficiaries per Physician	50
Physicians	500
Percentage of Quality Targets Met	50%
Average Physician Income	\$200,000
Foregone Medicare FFS Revenues as a Percentage of Medicare Savings	40%

SOURCE: Health Economics Research, Inc.

Note in particular that in this simulation we are assuming that the participating PGP saves 10% of expected per capita expenditures, i.e., the Medicare savings rate is 10%.¹⁷ Note also that we assume that 40% of the Medicare savings generated by the PGP represents foregone revenues. Foregone revenue is less than 100% because some of the reductions in volume are likely to occur among non-participating providers. For example, if a PGP does not have an affiliated hospital, a reduction in hospital admissions will not be foregone revenue to the PGP. The 40% assumption is derived from calculations with Medicare claims data of the proportion of services provided by PGPs and their affiliated providers to assigned beneficiaries (see Tompkins et al. 1996). We assume that this proportion equals the portion of Medicare savings that is foregone revenues, i.e., that under the demonstration the participating PGP proportionally reduces its own services and those of non-affiliated providers.

The parameter values for the number of assigned beneficiaries per physician, and the number of physicians, are, respectively, 50 and 500. This reflects simulations HER performed on Medicare 2000 Part B physician/supplier claims for nine PGPs. The number of assigned beneficiaries¹⁸ per physician ranged from 23 to 62, and the number of physicians ranged from 286 to 2048. The implied parameter value for number of assigned beneficiaries is thus 25,000 ($50 \times 500 = 25,000$).

¹⁷ Specifically, we assume that the growth rate in expenditures between the base and performance years for the PGP's market area is 6%, and the growth rate in expenditures between the base and performance years for beneficiaries assigned to the PGP is -4.6%, both for the PGP and its affiliated providers, and for non-affiliated providers. With an adjusted base year expenditures of \$6,500, this results in target expenditures of \$6,890 ($\$6,500 \times 1.06 = \$6,890$) and performance year expenditures of \$6,201 ($\$6,500 \times 0.954 = \$6,201$). Thus Medicare savings rate is $(\$6,890 - \$6,201) \div \$6,801 = 10\%$.

¹⁸ To determine the number of assigned beneficiaries for each PGP, we assumed that among the beneficiaries with at least one evaluation and management (E&M) visit at the PGP, that 67% received more E&M services from the PGP than from any other provider. See McCall *et al.*, 1998.

The parameter value for adjusted base year per capita Medicare FFS expenditures for assigned beneficiaries is \$6,500. In 1997, the per capita annualized Medicare FFS expenditures was \$5,157 (see Pope et al. 2000). We assume an average annual growth rate of 6% between 1997 and 2001, which results in per capita expenditures of approximately \$6,500 for 2001 ($\$5,157 \times 1.06^4 = \$6,511$). Although this is a somewhat arbitrary assumption, the growth rate in Medicare program payments per beneficiary served between 1990 and 1998 was 5.5% (HCFA Review 2001).

Any parameter assumption for the percentage of demonstration quality targets attained by the PGP is inherently arbitrary because there is no way to know in advance of the PGP demonstration how a participating PGP will respond to the bonus incentives provided by the demonstration. The parameter value for the percentage of quality targets attained is 50%. The parameter value for mean annual physician income is \$200,000. This is based on results from the Socioeconomic Monitoring System (SMS) survey (see Physician Socioeconomic Statistics 2000-2002 Edition). For 1998, the mean annual physician net income after expenses before taxes was \$194,000.

Bonus payments and revenue changes for the typical participating PGP are simulated based on sharing rates of 50%, 65%, 80%, and 95%. A sharing rate of 95% is simulated because Medicare historically discounted fee-for-service per capita expenditures by 5% in paying HMOs. A sharing rate of 50% is simulated because a natural default rate would be 50/50 sharing of savings between the PGP and the government. Steps for calculating bonus payments are outlined in Part I. The change in revenues for a PGP takes account of both bonus payments paid to the PGP and the FFS

revenues the PGP forgoes in generating Medicare savings. Bonus payments and revenue changes are simulated for various sharing rates, respectively, in Tables 4-2 and 4-3.

As shown in Table 4-2, bonus payments (post-withhold bonus) per participating PGP physician are about \$11,000, \$14,000, \$18,000, and \$21,000, for, respectively, sharing rates of 50%, 65%, 80%, and 95%. As a percentage of physician income, bonus payments are 5.5%, 7.1%, 8.8%, and 10.4%, and as a percentage of the PGP's projected Medicare FFS revenues for assigned beneficiaries absent the demonstration, they are 8.0%, 10.4%, 12.8%, and 15.1%.¹⁹ Based on bonus payments, a sharing rate of 50% would give a PGP sufficient incentive to participate in the demonstration.

However, a PGP is likely to be more interested in its change in Medicare revenues than its bonus payments. The revenue change subtracts from the PGP's bonus payments foregone FFS reimbursements from lower service volumes used to generate Medicare savings. As shown in Table 4-3, change in revenues pre-withhold per PGP physician are about \$900, \$5,300, \$9,600, and \$14,000, for, respectively, sharing rates of 50%, 65%, 80%, and 95%. As a percentage of physician income, change in revenues is 0.4%, 2.6%, 4.8%, and 7.0%, and as a percentage of the PGP's projected Medicare FFS revenues for assigned beneficiaries absent the demonstration they are 0.6%, 3.8%, 7.0%, and 10.2%. Based on change in revenues, a sharing rate of 50% would not give a PGP sufficient incentive to participate in the demonstration.

¹⁹ The PGP's projected Medicare FFS revenues for assigned beneficiaries absent the demonstration is based on the assumption that absent the demonstration, the growth rate of expenditures for assigned beneficiaries would equal the growth rate in expenditures for the PGP's market area, which is assumed to be 6%.

Table 4-2

Simulated Bonus Payments for Various Sharing Rates¹

	Sharing Rates			
	50%	65%	80%	95%
<u>Bonus Earned (Pre-Withhold Bonus)</u>				
Bonus	\$7,320,625	\$9,516,813	\$11,713,000	\$13,909,188
Bonus per Assigned Beneficiary	\$293	\$381	\$469	\$556
Bonus per Physician	\$14,641	\$19,034	\$23,426	\$27,818
Bonus as a % of Physician Income	7.3%	9.5%	11.7%	13.9%
Bonus as a % of PGP's Projected Medicare FFS Revenues for Assigned Beneficiaries Absent the Demonstration	10.6%	13.8%	17.0%	20.2%
<u>Withhold</u>				
Withhold Rate	25%	25%	25%	25%
Withhold	\$1,830,156	\$2,379,203	\$2,928,250	\$3,477,297
Withhold per Assigned Beneficiary	\$73	\$95	\$117	\$139
Withhold per Physician	\$3,660	\$4,758	\$5,857	\$6,955
Withhold as a % of Physician Income	1.8%	2.4%	2.9%	3.5%
Withhold as a % of PGP's Projected Medicare FFS Revenues for Assigned Beneficiaries Absent the Demonstration	2.7%	3.5%	4.3%	5.0%
<u>Bonus Payment (Post-Withhold Bonus)</u>				
Bonus	\$5,490,469	\$7,137,609	\$8,784,750	\$10,431,891
Bonus per Assigned Beneficiary	\$220	\$286	\$351	\$417
Bonus per Physician	\$10,981	\$14,275	\$17,570	\$20,864
Bonus as a % of Physician Income	5.5%	7.1%	8.8%	10.4%
Bonus as a % of PGP's Projected Medicare FFS Revenues for Assigned Beneficiaries Absent the Demonstration	8.0%	10.4%	12.8%	15.1%

¹See Table 4-1 for parameter values assumed for this simulation.

Table 4-3

Simulations of PGP Change in Medicare Revenues for Various Sharing Rates¹

	Sharing Rates			
	50%	65%	80%	95%
<u>Change in Revenues Pre-Withhold</u>				
Change in Revenues	\$430,625	\$2,626,813	\$4,823,000	\$7,019,188
Change in Revenues per Assigned Beneficiary	\$17	\$105	\$193	\$281
Change in Revenues per Physician	\$861	\$5,254	\$9,646	\$14,038
Change in Revenues as a % of Physician Income	0.4%	2.6%	4.8%	7.0%
Change in Revenues as a % of PGP's Projected Medicare FFS Revenues for Assigned Beneficiaries Absent the Demonstration	0.6%	3.8%	7.0%	10.2%
<u>Withhold</u>				
Withhold Rate	25%	25%	25%	25%
Withhold	\$1,830,156	\$2,379,203	\$2,928,250	\$3,477,297
Withhold per Assigned Beneficiary	\$73	\$95	\$117	\$139
Withhold per Physician	\$3,660	\$4,758	\$5,857	\$6,955
Withhold as a % of Physician Income	1.8%	2.4%	2.9%	3.5%
Withhold as a % of PGP's Projected Medicare FFS Revenues for Assigned Beneficiaries Absent the Demonstration	2.7%	3.5%	4.3%	5.0%
<u>Change in Revenues Post-Withhold</u>				
Change in Revenues	-\$1,399,531	\$247,609	\$1,894,750	\$3,541,891
Change in Revenues per Assigned Beneficiary	-\$56	\$10	\$76	\$142
Change in Revenues per Physician	-\$2,799	\$495	\$3,790	\$7,084
Change in Revenues as a % of Physician Income	-1.4%	0.2%	1.9%	3.5%
Change in Revenues as a % of PGP's Projected Medicare FFS Revenues for Assigned Beneficiaries Absent the Demonstration	-2.0%	0.4%	2.8%	5.1%

¹See Table 4-1 for parameter values assumed for this simulation.

Although for a sharing rate of 65%, the change in revenues pre-withhold per physician and as a percentage of physician income are \$5,254 and 2.6%, post-withhold they are only \$495 and 0.2%. Thus a sharing rate of 65% might not give a PGP sufficient incentive to participate in the demonstration. However, for a sharing rate of 80%, the change in revenues pre-withhold per physician and as a percentage of physician income are \$9,600 and 4.8%, and post-withhold they are \$3,790 and 1.9%. HER's recommendation is to use a sharing rate of 80%. It gives the PGP a sufficient incentive to participate in the demonstration, and yet allows for significant Medicare program savings under the demonstration.

Recommendation 5: Bonus Payments, Settlements, and Withdrawal

- Bonuses may be earned by participating PGPs in performance years in which the organization generates Medicare Savings.
- Medicare Losses accrue to participating PGPs in performance years in which PGP expenditures exceed their Target.
- A PGP's accrued Medicare Loss carried forward from the prior performance year, if any, is deducted from its Medicare Savings before bonuses are determined.
- The annual PGP bonus, if any, will be paid at annual settlement, with a portion withheld until final demonstration settlement contingent on future performance.
- The maximum bonus that can be earned by a PGP in a year (bonus payments plus withheld amounts) is limited to 15% of target Medicare expenditures for beneficiaries assigned to that organization in that year.
- At final demonstration settlement, CMS will remit withheld bonus amounts to the PGP. Accrued losses will be deducted from the amount returned by CMS to the PGP. Even if accrued losses exceed withheld bonuses, at most the PGP will forfeit withheld bonuses at final settlement.
- If a participating PGP withdraws from the demonstration before its completion, it will forfeit all withheld bonus payments.

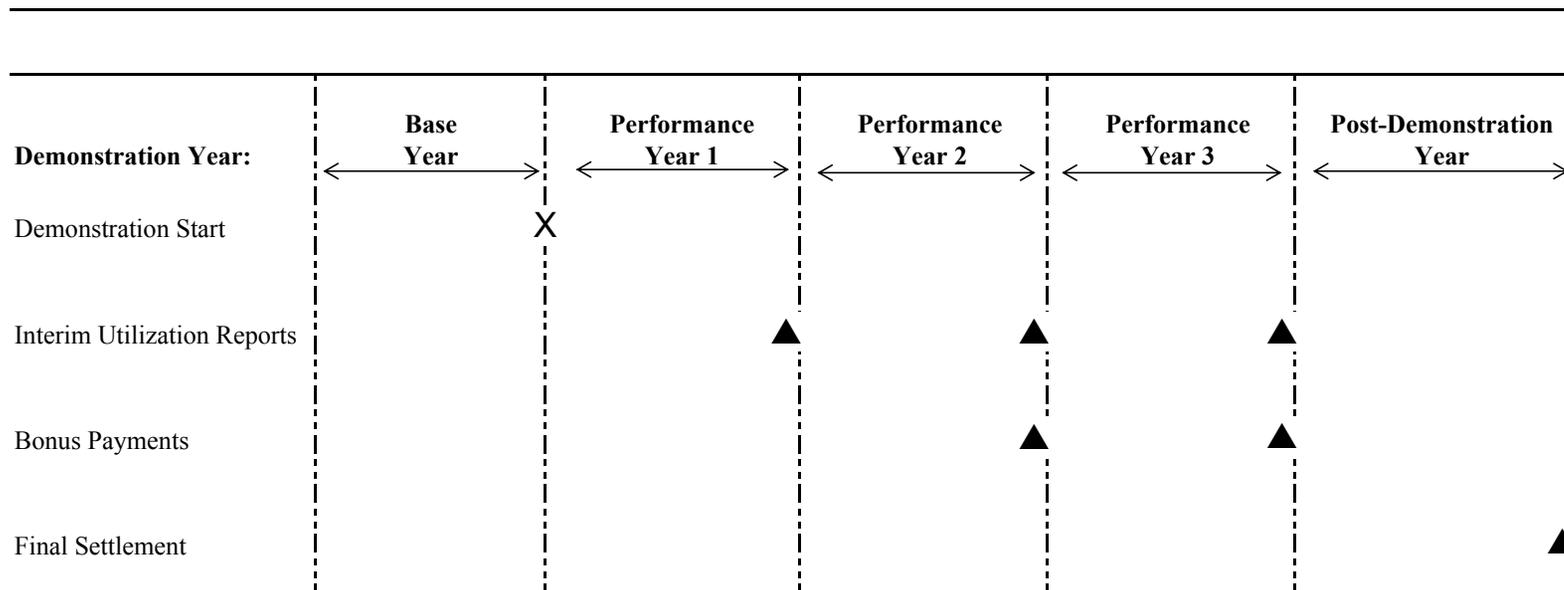
Figure 5-1 shows the demonstration schedule. The demonstration will begin with performance year one (PY1). An interim report on utilization in PY1 versus the base year will be provided to the participating PGP at the end of PY1 (see Recommendation 9). Medicare claims data for PY1 will be considered complete six months after the end of the year. Claims data for the PGP and its comparison group will be obtained from Medicare datafiles and processed by the demonstration technical support contractor over the next 6 months. An annual bonus settlement cycle will occur approximately one year after the end of PY1, and then one year after PY2 and one year after PY3. Final demonstration settlement will occur approximately one year after the end of PY3, simultaneously with the PY3 annual bonus settlement.

Bonus payments may be made to a PGP if it generates Medicare savings in a performance year. A “Medicare Loss” will accrue in any performance year in which PGP per capita expenditures exceed its Target.²⁰ The PGP’s bonus pool in a performance year is based on the PGP’s Medicare Savings or Loss for that year combined with its accrued Medicare Losses, if any, carried forward from the prior performance year. If this combined amount is greater than zero, a portion of the amount will be paid to the PGP and the remainder will be withheld contingent on future performance. (Withheld amounts will be returned to the PGP as part of final settlement.) To avoid incentives for excessive cost cutting, the maximum bonus that can be earned by a PGP in a year (bonus payments plus withheld amounts) is limited to 15% of target Medicare reimbursements for assigned beneficiaries. Participating organizations would simply forgo bonuses in

²⁰ Small annual Medicare Savings or Losses, that may be due to chance, will be considered to be zero. See Recommendation 7.

Figure 5-1

PGP Demonstration Timeline



SOURCE: Health Economics Research, Inc.

excess of this limit, i.e., bonuses in excess of the 15% limit will not be paid, withheld, or accrued for payment in future years.

An example, as shown in Table 5-1, will help clarify these concepts and steps. We simplify the example by assuming that a participating organization meets all quality targets each year of the demonstration so that the PGP earns all of its bonus pool. Suppose a participating organization generates positive Medicare Savings of \$3,900,000 in performance year 1 (PY1), a Medicare Loss of -\$6,030,000 in PY2, and Medicare savings of \$4,350,000 in PY3. After applying the Sharing Ratio of 80%, the PGP's bonus pool for PY1 is \$3,120,000. Assuming a withholding rate of 25% (see Recommendation 6), \$2,340,000 is paid to the PGP as the PY1 bonus and \$780,000 is withheld for final settlement. In PY2, Medicare Savings are negative. A Loss of -\$6,030,000 is accrued, and is carried forward to PY3. In PY3, Medicare savings are \$4,350,000. But when the accrued Loss of -\$6,030,000 carried forward from the previous year is charged against the Annual Savings, the resulting amount for PY3 is -\$1,680,000 and no PY3 bonus is paid.

Final demonstration settlement will occur approximately one year after the end of PY3. At settlement, the withheld annual bonus payments will be debited by 80% of the Accrued Loss, if any, at the end of the demonstration.²¹ Continuing the example of the preceding paragraph, withheld bonuses total \$780,000. Eighty percent of the Accrued Loss at the end of the demonstration is -\$1,344,000 (80% of \$1,680,000). The sum is -\$564,000. Therefore, the PGP forfeits the \$780,000 in withheld bonuses, i.e., CMS does

²¹ The Accrued Loss is adjusted by the sharing ratio, 80%. Just as Medicare Savings are shared between the PGP and CMS, so too losses are shared.

Table 5-1**Example of PGP Demonstration Bonus Computation**

	Performance <u>Year 1</u>	Performance <u>Year 2</u>	Performance <u>Year 3</u>	Final <u>Settlement</u>
Annual Medicare Savings or Loss	\$3,900,000	-\$6,030,000	\$4,350,000	--
Accrued Medicare Loss from Previous Year	0	0	-6,030,000	-1,680,000
Sum ¹	3,900,000	-6,030,000	-1,680,000	--
Annual Bonus Pool ²	3,120,000	0	0	--
Bonus Paid ³	2,340,000	0	0	--
Bonus Withheld Until Final Settlement ⁴	780,000	0	0	780,000
Final Settlement Amount ⁵	--	--	--	0

¹Sum of Annual Medicare Savings/Loss and Accrued Medicare Loss from previous year.
If negative, carried forward to next year.

²80% (the sharing ratio) of the sum of annual accrued Medicare Savings/Losses if greater than zero, zero otherwise.

³75% (100% - the withholding %) of the bonus pool.

⁴25% (the withhold %) of the bonus pool.

⁵Sum of 80% of accrued Medicare Loss from Year 3, if any, and bonus withheld until final settlement.
In this example, this sum equal -564,000. However, the participating PGP never loses more than
the withheld bonus amounts at final settlement. So the final settlement amount is zero in this example.

NOTE: Assumes that all quality targets are met each year, a sharing ratio of 80%, and a bonus
withholding rate of 25%.

SOURCE: Health Economics Research, Inc.

not return the withheld bonuses at demonstration settlement. The participating organization is never required to return to CMS at settlement any paid bonuses it has received during the demonstration.

The PGP demonstration is intended to measure longer-run, permanent changes in PGP cost and quality behavior, not transitory year-to-year fluctuations. For this reason, if a participating organization withdraws from the demonstration prior to its full, three-year completion, it will forfeit all withheld bonus payments. This policy avoids incentives for the PGP to drop out when it earns large bonuses in PY1 and PY2, but foresees a decline in its performance in PY3 (or simply does not want to chance a decline in performance in PY3) that would be charged against its earlier withheld bonuses.

Simulations of Per Capita Bonuses Under Alternative Expenditure Growth Rate Scenarios. It is useful to gain further insight into the possible multi-year patterns of PGP bonuses and settlements that may occur under the demonstration. To do so, we simulated bonus payments and settlement amounts under alternative expenditure growth rate scenarios. These results are presented in Tables 5-2 and 5-3. Table 5-2 presents the full simulation results. Table 5-3 is a summary of PGP cash flow (annual bonus payments and settlement amounts) under alternative scenarios. These simulations are done on a per capita basis because it is easier to compare relative magnitudes if everything is on a per capita basis. Results can be scaled up to total dollars by multiplying all quantities by an assumed number of assigned PGP beneficiaries (e.g., 30,000). By adding the per capita amounts, we are making the simplifying assumption that the number of assigned beneficiaries is the same in each performance year. We also

Table 5-2

Simulations of Per Capita Bonus Payments Under Alternative Expenditure Growth Rate Scenarios

Year	Expenditure Growth Rate		Per Capita Expenditures		Medicare Savings or Loss	Annual Bonus Pool	Accrued Loss From Prev Year	Sum of Annual, Accrued Bonus	Bonus Paid	Bonus Withheld	Accrued Loss	Final Settlement	Summary
	PGP Site	Comparison Group	PGP Site	Target						Until Final Settlement	Carried Forward	Amount	Totals
Scenario 1: Good Performance All Years													
BY	--	--	\$6,500	--	--	--	--	--	--	--	--	--	--
PY1	3%	8%	6,695	\$7,020	325	\$260	\$0	\$260	\$195	\$65	\$0	--	--
PY2	3	8	6,896	7,582	686	549	0	549	411	137	0	--	--
PY3	3	8	7,103	8,188	1085	868	0	868	651	217	0	--	--
Settlement Amount							0			419		419	
Cash Flow at Settlement (Settlement+PY3 Bonus)									651			419	1,070
Total Final Net Payments to PGP									1,258			419	1,677
Cumulative Amounts Three Performance Years					2096	1,677			1,258	419	0		
Scenario 2: Good Performance First Year, Poor Performance Second Year, Good Performance Third Year													
BY	--	--	6,500	--	--	--	--	--	--	--	--	--	--
PY1	3	8	6,695	7,020	325	260	0	260	195	65	0	--	--
PY2	8	3	7,231	7,231	0	0	0	0	0	0	0	--	--
PY3	3	8	7,448	7,809	362	289	0	289	217	72	0	--	--
Settlement Amount							0			137		137	
Cash Flow at Settlement (Settlement+PY3 Bonus)									217			137	354
Total Final Net Payments to PGP									412			137	549
Cumulative Amounts Three Performance Years					687	549			412	137	0		

Table 5-2 (Continued)

Simulations of Per Capita Bonus Payments Under Alternative Expenditure Growth Rate Scenarios

Year	Expenditure Growth Rate		Per Capita Expenditures		Medicare Savings or Loss	Annual Bonus Pool	Accrued Loss From Prev Year	Sum of Annual, Accrued Bonus	Bonus Paid	Bonus Withheld	Accrued Loss	Final Settlement	Summary
	PGP Site	Comparison Group	PGP Site	Target						Until Final Settlement	Carried Forward	Amount	Totals
Scenario 3: Poor Performance First Year, Good Performance Years 2 and 3													
BY	--	--	\$6,500	--	--	--	--	--	--	--	--	--	--
PY1	8%	3%	7,020	\$6,695	-\$325	-\$260	\$0	-\$260	\$0	\$0	-\$260	--	--
PY2	3	8	7,231	7,231	0	0	-260	-260	0	0	-260	--	--
PY3	3	8	7,448	7,809	362	289	-260	29	22	7	0	--	--
Settlement Amount							0			7		7	
Cash Flow at Settlement (Settlement+PY3 Bonus)									22			7	29
Total Final Net Payments to PGP									22			7	29
Cumulative Amounts Three Performance Years					37	29			22	7	0		
Scenario 4: Good Performance First Year, Neutral Performance Second Year, Very Poor Performance Third Year													
BY	--	--	6,500	--	--	--	--	--	--	--	--	--	--
PY1	5	8	6,825	7,020	195	156	0	156	117	39	0	--	--
PY2	8	8	7,371	7,582	211	168	0	168	126	42	0	--	--
PY3	12	3	8,256	7,809	-446	-357	0	-357	0	0	-357	--	--
Settlement Amount							-357			81		0	
Cash Flow at Settlement (Settlement+PY3 Bonus)									0			0	0
Total Final Net Payments to PGP									243			0	243
Cumulative Amounts Three Performance Years					-41	-33			243	81	-357		

Table 5-2 (Continued)

Simulations of Per Capita Bonus Payments Under Alternative Expenditure Growth Rate Scenarios

Year	Expenditure Growth Rate		Per Capita Expenditures		Medicare Savings or Loss	Annual Bonus Pool	Accrued Loss From Prev Year	Sum of Annual, Accrued Bonus	Bonus Paid	Bonus Withheld	Accrued Loss	Final Settlement Amount	Summary Totals
	PGP Site	Comparison Group	PGP Site	Target						Until Final Settlement	Carried Forward		
Scenario 5: Good Performance First Year, Poor Performance Second Year, Good Performance Third Year													
BY	--	--	\$6,500	--	--	--	--	--	--	--	--	--	--
PY1	3%	5%	6,695	\$6,825	\$130	\$104	\$0	\$104	\$78	\$26	\$0	--	--
PY2	8	3	7,231	7,030	-201	-161	0	-161	0	0	-161	--	--
PY3	3	8	7,448	7,592	145	116	-161	-45	0	0	-45	--	--
Settlement Amount							-45			26		0	
Cash Flow at Settlement (Settlement+PY3 Bonus)									0			0	0
Total Final Net Payments to PGP									78			0	78
Cumulative Amounts Three Performance Years					74	59			78	26	-45		
Scenario 6: Neutral Performance First Year, Good Performance Second and Third Years													
BY	--	--	6,500	--	--	--	--	--	--	--	--	--	--
PY1	3	3	6,695	6,695	0	0	0	0	0	0	0	--	--
PY2	3	8	6,896	7,231	335	268	0	268	201	67	0	--	--
PY3	3	8	7,103	7,809	706	565	0	565	424	141	0	--	--
Settlement Amount							0			208		208	
Cash Flow at Settlement (Settlement+PY3 Bonus)									424			208	632
Total Final Net Payments to PGP									625			208	833
Cumulative Amounts Three Performance Years					1041	833			625	208	0		

Table 5-2 (Continued)

Simulations of Per Capita Bonus Payments Under Alternative Expenditure Growth Rate Scenarios

Year	Expenditure Growth Rate		Per Capita Expenditures		Medicare Savings or Loss	Annual Bonus Pool	Accrued Loss From Prev Year	Sum of Annual, Accrued Bonus	Bonus Paid	Bonus Withheld	Accrued Loss	Final Settlement	Summary
	PGP Site	Comparison Group	PGP Site	Target						Until Final Settlement	Carried Forward	Amount	Totals
Scenario 7: Poor Performance First and Second Years, Good Performance Third Year													
BY	--	--	6,500	--	--	--	--	--	--	--	--	--	--
PY1	8	3	7,020	6,695	-325	-260	0	-260	0	0	-260	--	--
PY2	8	3	7,582	6,896	-686	-549	-260	-809	0	0	-809	--	--
PY3	3	8	7,809	7,448	-362	-289	-809	-1098	0	0	-1,098	--	--
Settlement Amount							-1098			0		0	
Cash Flow at Settlement (Settlement+PY3 Bonus)									0			0	0
Total Final Net Payments to PGP									0			0	0
Cumulative Amounts Three Performance Years					-1372	-1098			0	0	-1,098		
Scenario 8: Good Performance First Year, Neutral Performance Second Year, and Poor Performance Third Year													
BY	--	--	\$6,500	--	--	--	--	--	--	--	--	--	--
PY1	5%	8%	6,825	\$7,020	\$195	\$156	0	\$156	\$117	\$39	\$0	--	--
PY2	8	8	7,371	7,582	211	168	0	168	126	42	0	--	--
PY3	8	3	7,961	7,809	-152	-121	0	-121	0	0	-121	--	--
Settlement Amount							-121			81		0	
Cash Flow at Settlement (Settlement+PY3 Bonus)									0			0	0
Total Final Net Payments to PGP									243			0	243
Cumulative Amounts Three Performance Years					254	203			243	81	-121		

Table 5-2 (Continued)

Simulations of Per Capita Bonus Payments Under Alternative Expenditure Growth Rate Scenarios

Year	Expenditure Growth Rate		Per Capita Expenditures		Medicare Savings or Loss	Annual Bonus Pool	Accrued Loss From Prev Year	Sum of Annual, Accrued Bonus	Bonus Paid	Bonus Withheld	Accrued Loss Carried Forward	Final Settlement Amount	Summary Totals
	PGP Site	Comparison Group	PGP Site	Target						Until Final Settlement			
Scenario 9: Poor Performance First Year, Good Performance Second and Third Years													
BY	--	--	6,500	--	--	--	--	--	--	--	--	--	--
PY1	5	3	6,825	6,695	-130	-104	0	-104	0	0	-104	--	--
PY2	3	8	7,030	7,231	201	161	-104	57	43	14	0	--	--
PY3	3	8	7,241	7,809	568	455	0	455	341	114	0	--	--
Settlement Amount							0			128		128	
Cash Flow at Settlement (Settlement+PY3 Bonus)									341			128	469
Total Final Net Payments to PGP									384			128	511
Cumulative Amounts Three Performance Years					639	511			384	128	0		

Assumes:

- 1) All quality targets are met.
- 2) Baseline per capita PGP expenditures are \$6,500.
- 3) Sharing ratio = 0.80.
- 4) A 25% annual bonus withholding rate; withheld amounts are returned at settlement.
- 5) Ignores interest payments, time value of money, present value.
- 6) The same number of assigned beneficiaries in each performance year.

SOURCE: Health Economics Research, Inc.

Table 5-3

**Maximum Per Capita PGP Bonus Payments by Performance Year (PY)
Under Alternative Expenditure Growth Scenarios
(Summary of Table 5-2)**

	Per Capita Expenditure Growth Rate		Medicare Savings Rate	Annual Bonus Paid
	<u>PGP</u>	<u>Comparison</u>		
Scenario 1: Good Performance All Three Years				
PY1	3%	8%	5%	\$195
PY2	3	8	5	411
PY3	3	8	5	651
Final Settlement				419
Total final net payments to PGP				1,677
Scenario 2: Good Performance First Year, Poor Performance Second Year, Good Performance Third Year				
PY1	3	8	5	195
PY2	8	3	-5	0
PY3	3	8	5	217
Final Settlement				137
Total final net payments to PGP				549
Scenario 3: Poor Performance First Year, Good Performance Years 2 and 3				
PY1	8	3	-5	0
PY2	3	8	5	0
PY3	3	8	5	22
Final Settlement				7
Total final net payments to PGP				29
Scenario 4: Good Performance First Year, Neutral Performance Second Year, Very Poor Performance Third Year				
PY1	5	8	3	117
PY2	8	8	0	126
PY3	12	3	-9	0
Final Settlement				0
Total final net payments to PGP				243
Scenario 5: Good Performance First Year, Poor Performance Second Year, Good Performance Third Year				
PY1	3	5	2	78
PY2	8	3	-5	0
PY3	3	8	5	0
Final Settlement				0
Total final net payments to PGP				78

Table 5-3 (Continued)

**Maximum Per Capita PGP Bonus Payments by Performance Year (PY)
Under Alternative Expenditure Growth Scenarios
(Summary of Table 5-2)**

	Per Capita Expenditure Growth Rate		Medicare Savings Rate	Annual Bonus Paid
	<u>PGP</u>	<u>Comparison</u>		
Scenario 6: Neutral Performance First Year, Good Performance Second and Third Years				
PY1	3%	3%	0%	\$0
PY2	3	8	5	201
PY3	3	8	5	424
Final Settlement				208
Total final net payments to PGP				833
Scenario 7: Poor Performance First and Second Years, Good Performance Third Year				
PY1	8	3	-5	0
PY2	8	3	-5	0
PY3	3	8	5	0
Final Settlement				0
Total final net payments to PGP				0
Scenario 8: Good Performance First Year, Neutral Performance Second Year, and Poor Performance Third Year				
PY1	5	8	3	117
PY2	8	8	0	126
PY3	8	3	-5	0
Final Settlement				0
Total final net payments to PGP				243
Scenario 9: Poor Performance First Year, Good Performance Second and Third Years				
PY1	5	3	-2	0
PY2	3	8	5	43
PY3	3	8	5	341
Final Settlement				128
Total final net payments to PGP				511

Assumes:

- 1) All quality targets are met.
- 2) Baseline per capita PGP expenditures are \$6,500.
- 3) Sharing ratio = 0.80.
- 4) A 25% annual bonus withholding rate; withheld amounts are returned at settlement.
- 5) Ignores interest payments, time value of money, present value.
- 6) Assumes the same number of assigned beneficiaries in each performance year.

SOURCE: Health Economics Research, Inc.

assume that all quality targets are met in each year, so the amounts represent maximum bonuses. A sharing ratio of 80% and a bonus withholding rate of 25% are assumed. Also, for simplicity, the bonus payment threshold is assumed to be zero (see Recommendation 7).²² This permits us to focus on the bonus payment and accrual mechanisms as well as the final settlement process. We ignore the time value of money (interest payments, present values). Finally, we note that the “accounting” is done a little differently in Table 5-2 than in Table 5-1. In Table 5-1 accrued Medicare losses are combined with Medicare savings before the annual bonus pool is determined. In Table 5-2, losses are accrued and carried forward as part of the bonus pool. These two methods of accounting are equivalent, i.e., they result in the same bonus payments, under the assumptions made for these simulations.

The first scenario assumes good PGP cost saving performance in each of the three demonstration performance years, a 5% savings versus the comparison group in each year. Under this scenario, the PGP receives a bonus for every performance year, plus an additional payment at settlement returning withheld bonus amounts. There are no accrued losses carried forward from year to year. The bonus amounts paid rise every year because of the cumulative effect of a lower PGP than comparison group expenditure growth rate. As in most of the simulations, the total final net payments to the PGP (the sum of bonus payments and the settlement amount) equal the sum of the annual bonus pools (i.e., the PGP’s portion of cumulative Medicare savings).

²² Recommendation 7 concerns the establishment of a threshold rate for the payment of bonuses and the accruals of losses.

The second scenario is of good cost control performance by the PGP in the first year, poor performance in the second year, and good performance in the third year. In this "mixed" situation, bonuses are paid in the first and third years, but not in the second. There are no accrued losses and the settlement amount is again positive (a return of the bonus withholds).

The third scenario is of poor PGP performance in the first year, then good performance in the second and third years. In this scenario, a loss is accrued in the first year, and is carried forward to the third year since the PY2 annual bonus pool is zero. The loss carryforward from PY1 mostly offsets the annual bonus earned in PY3. On net, the PGP earns a small bonus in PY3 in this scenario.

It is instructive to compare the second and third scenarios. In both cases, PY3 PGP per capita expenditures is identical, \$7,448 (comparison group expenditures are the same in each year in the two scenarios). However, in Scenario 2, the PGP earns final net payments of \$549 per capita from the demonstration, whereas in Scenario 3, final net payments per capita are only \$29. At first glance it may be surprising that bonus payments are so different when final PGP per capita expenditures are the same. This comparison illustrates that PGP Demonstration bonus payments are path dependent. That is, total bonus payments to the PGP depend on the timing of its cost control performance as well as where its costs end up. Comparing two PGPs with equal average expenditure growth rates over the three demonstration years, the PGP that achieves its cost savings in earlier years will earn much larger bonuses than the PGP that only controls its cost late in the performance period. This is appropriate, because total Medicare Savings, which equal cost level multiplied by years elapsed, are greater in the former scenario than in the

latter. In other words, the PGP Demonstration is not a test of average rate of cost growth over the three year demonstration period, but rather a test of total Medicare expenditures saved over this period. It is possible that a PGP with a higher final per capita expenditure level in PY3 could earn larger total bonuses under the demonstration than a PGP with a lower final expenditure level if the former PGP controlled costs better earlier in the demonstration.

Scenario 4 shows a case of good performance in the first year, neutral performance in the second year, and poor performance in the third year. In this situation, a bonus is paid in the first two years, but a large loss is accrued in the third year. At settlement, the third year loss (-\$357) is combined with the withheld bonuses (\$81) from PY1 and PY2 to equal -\$276. Although the final settlement amount is negative, the PGP will keep the paid bonuses of \$243 (per assigned beneficiary). The PGP forfeits the \$81 in withheld bonuses but is not required to make further remittances to CMS to offset the negative final settlement amount. So the settlement amount in Table 5-2 for Scenario 4 is shown as zero.

Scenario 5 is similar to Scenario 2 in its pattern of good, bad, then good performance. But in Scenario 5, performance is not as good in PY1 (a 2% Medicare Savings rate versus a 5% rate in Scenario 2). For this reason, the accrued bonus at the end of the demonstration is negative, even when offset by the (positive) amount withheld from the PY1 bonus. Thus, unlike Scenario 2, the PGP forfeits the withheld bonus.

Scenario 6 presents a situation of neutral performance in the first year, then good performance in the second and third years. Except for PY1, bonuses are paid. At settlement the withheld bonuses are returned to the PGP. This scenario shows that a PGP can still do well under the demonstration even if it is "not able to get its act together" quickly and only controls costs after the first demonstration year.

Scenario 7 shows poor performance in the first two demonstration years and good performance only in the last year. In this case, the PGP never "gets out of the hole". Its annual and accrued bonus pools are never positive, and it is never paid a bonus. But even though the accrued bonus is quite negative at the end of the demonstration, the PGP owes CMS nothing at settlement. The PGP does not lose anything from the demonstration in spite of its poor performance.

Scenario 8 consists of good performance the first year, neutral performance the second year, and poor performance the third year. The PGP earns a bonus in the first year. It also earns a PY2 bonus despite its neutral performance in that year. This is because its good first year performance carries over to the second year, and cumulatively, its expenditure growth rate remains below its comparison group. In the third year, the PGP's poor performance drives its annual bonus negative, and the PGP forfeits the \$81 in withheld bonuses at settlement. But on net the PGP earns \$243 per capita from the demonstration, the same outcome as Scenario 4. This scenario again demonstrates that good performance early in the demonstration can drive attractive bonus payments because the early cost control is carried forward into later years.

Scenario 9 is similar to Scenario 3 in that performance is poor in PY1, but good in PY2 and PY3. However, in Scenario 9, performance is not as bad (relative to the comparison group) in PY1 as in Scenario 3 (-2% Medicare Savings rate versus -5%). Scenario 9 demonstrates that even with somewhat poor performance in PY1, the PGP can still "climb out of the hole" and earn a reasonable bonus by performing well in PY2 and PY3. The final net per capita payment to the PGP in Scenario 9 is \$511 versus only \$29 in Scenario 3. In short, unless the PGP's performance in PY1 is quite bad as in Scenario 3, there is no reason for the PGP to become "discouraged" by a mediocre first year performance and lose hope of earning significant bonuses under the demonstration.

Appendix to Recommendation 5: Equations Behind Simulations

This appendix summarizes the equations used in the simulation model whose results are presented in Tables 5-2 and 5-3. These equations may be useful in understanding the simulations, in performing future simulations, and in developing a complete payment simulation model for the demonstration. We use the abbreviation "PCE" in the equations to denote "per capita expenditures". The function "Max(x;y)" returns the maximum of x and y.

$$\text{PY1 PGP PCE} = (\text{BY PGP PCE}) * (1 + \text{PY1 PGP PCE growth rate}).$$

$$\text{PY2 PGP PCE} = (\text{PY1 PGP PCE}) * (1 + \text{PY2 PGP PCE growth rate}).$$

$$\text{PY3 PGP PCE} = (\text{PY2 PGP PCE}) * (1 + \text{PY3 PGP PCE growth rate}).$$

$$\text{PY1 target PCE} = (\text{BY PGP PCE}) * (1 + \text{PY1 comparison group PCE growth rate}).$$

$$\text{PY2 target PCE} = (\text{PY1 target PCE}) * (1 + \text{PY2 comparison group PCE growth rate}).$$

$PY3 \text{ target PCE} = (PY2 \text{ target PCE}) * (1 + PY3 \text{ comparison group PCE growth rate}).$

$Medicare \text{ Savings} = \text{target PCE} - \text{PGP PCE}.$

$\text{Annual bonus pool} = (\text{sharing rate}) * (\text{Medicare Savings}).$

Accrued bonus or penalty from previous year = 0 in PY1; = accrued bonus or penalty carried forward from the previous year (see below) for all other years.

Sum of annual and accrued bonus pools = annual bonus pool + (accrued bonus or penalty from previous year).

$\text{Bonus paid} = (1 - \text{withholding rate}) * \text{Max}(\text{sum of annual and accrued bonus}; 0).$

$\text{Bonus withheld until final settlement} = \text{Max}(\text{sum of annual and accrued bonus} - \text{bonus paid}; 0).$

$\text{Accrued bonus or penalty carried forward} = (\text{sum of annual and accrued bonus}) - (\text{bonus paid}) - (\text{bonus withheld until final settlement}).$

$\text{Final settlement amount} = \text{Max}(\text{accrued bonus or penalty carried forward from PY3} + \text{sum of bonus amounts withheld in PY1, PY2, and PY3}; 0).$

Recommendation 6: Withholds

- HER recommends a 25 percent withholding rate on bonus payments.
- At the end of the demonstration, cumulative withheld amounts in excess of accrued losses will be returned to the PGP.

Background: When positive Medicare savings occur and a participating PGP is eligible for a bonus payment, should a portion of the bonus payment be withheld to offset possible losses in future years? The withhold issue is relevant when the PGP demonstration is viewed in a multi-year instead of a single-year context.

Over time, a participating PGP might accrue bonuses in some years (perhaps due to chance) and negative bonuses (losses) in other years. The question is whether full bonuses should be paid in the year they are accrued, or whether some portion should be withheld to offset future losses.

Discussion: The main argument against withholds is that participating PGPs might need the demonstration bonus funds to offset the lower Medicare FFS revenues that generated the bonus. This may be especially relevant for PGPs with affiliated hospitals that generated their bonus by reducing hospital admissions, which generally involve larger revenues than other types of medical services. Conversely, this argument may be less forceful if the participant does not have an affiliated hospital.

One argument in favor of withholding is that it is administratively more feasible than CMS having to *ex post* recover payments from participating PGPs. This is especially true if the magnitude of payments that need to be recovered are relatively small. Withholds also provide a measure of protection to the Medicare program in the event that a participating PGP that had received bonus payments decides to leave the demonstration prior to its scheduled end. That is, if it is difficult for Medicare to recover bonus payments, withholds ensure that Medicare can recover some of the positive bonuses.

A flat withholding rate is simpler to understand than a withholding rate schedule. A flat rate that is too high (e.g., 70 percent) could cause participating PGPs to have cashflow problems and, hence, drop out of the demonstration. A flat rate that is too low (e.g., 10 percent) is more likely to result in CMS having to recover payments from participating PGPs—a process that might be administratively difficult and politically

awkward.

To ascertain the financial (cash-flow) implications of different withholding rates, we simulated net Medicare revenues with a 25 percent withhold rate and a 50 percent withhold rate. The simulation was based on the following assumptions: 25,000 assigned FFS Medicare patients, 200 PGP physicians, an 80 percent sharing rate, a bonus payment threshold of two percent, a 5 percent Medicare savings rate, 40% of Medicare savings are foregone revenues to the participating PGP, and 50 percent attainment of quality targets. Net Medicare revenues take into account the foregone revenues necessary to generate Medicare savings as well as the PGP demonstration bonus payments.

With a 25 percent withhold rate, the PGP's net Medicare revenues increase by \$893,750 or \$4,469 per physician (Table 6-1). With a 50 percent withhold rate, the PGP's Medicare revenues fall \$487,500 or \$2,438 per physician. The increased total bonus of \$1,381,250 (\$4,143,750 minus \$2,762,500) paid out to the PGP, instead of being withheld by withheld by CMS, is responsible for the increase in total net Medicare revenues.

The simulation shows that a 50 percent withholding rate could result in PGP revenue loss under the demonstration net of withhold and, possibly, cash flow problems. Thus, we recommend a 25 percent withholding rate for the demonstration. As for possible problems in recovering previous bonus payments in the event of future cumulative losses, in addition to the 25 percent withhold, Medicare is financially protected by:

Table 6-1
Cash Flow by Bonus Pool and Withholding Rate

	Withholding Rate	
	<u>25%</u>	<u>50%</u>
<i>Changes in net Medicare Revenues</i>		
Total Change	\$893,750	-\$487,500
Total Change per Physician	4,469	-2,438
<i>Bonuses Paid-Out at Annual Settlement per Physician</i>		
Cost-Savings	17,063	11,375
Quality	3,656	2,438
Total	20,719	13,813

NOTE: Change in net Medicare revenues include both foregone revenues used to generate Medicare savings and bonus payments under the PGP demonstration.

SOURCE: Health Economics Research, Inc.

- guaranteed minimum savings of 20 percent of Medicare savings;
- the 2 percent annual threshold required for any bonus payments (see Recommendation 7), and
- the likelihood that participating PGPs will not achieve all their quality targets, generating additional Medicare program savings.

With regard to quality, because of the disparate quality indicators, each of which can be satisfied by either exceeding a threshold or demonstrating improvement, there is no simple combined index or metric by which to measure cumulative quality performance. Therefore, we recommend that no attempt be made to recover previous quality bonus payments in the event of the deterioration of quality performance. All quality bonuses are at risk under the demonstration if no cumulative cost savings are demonstrated. That is, at the end of the demonstration, no net quality bonuses will be

paid to participating PGPs if they have not generated Medicare (cost) savings. All amounts withheld from the quality portion of the PGP's annual bonus are at risk to cover future shortfalls in Medicare savings.

Given the brevity of the PGP demonstration, three years, we recommend that annual withhold amounts (in excess of negative bonus payments) be returned at the end of demonstration rather than being disbursed at interim periods prior to the end of the demonstration. Retaining the annual withhold amounts until the end of the demonstration provides Medicare yet another measure of financial protection against early withdrawal from the PGP demonstration.

Recommendation 7: Required Number of Beneficiaries for Participating Physician Group Practices and Comparison Groups, and Thresholds for Bonus Payment

- The PGP demonstration aims to pay bonuses for achieving real cost efficiencies, not for random fluctuations in costs. Two means of avoiding paying bonuses for random fluctuations are to require participating physician group practices (PGPs) and their comparison groups to have a sufficiently large number of assigned beneficiaries, and to only pay bonuses when cost savings exceed a threshold percentage amount. Larger numbers of beneficiaries improve the accuracy of cost estimation and a bonus threshold avoids paying a bonus for small cost differences that could be due to chance. HER recommends that participating PGPs and their comparison groups are assigned at least 15,000 to 20,000 beneficiaries. In addition, HER recommends a bonus threshold of 2.0%. This means that a bonus would not be paid unless the difference in the site and market expenditure growth rates exceeds 2.0%. However, if the threshold is met, the full bonus, not just the amount above 2%, would be paid. Medicare losses will only be accrued when the losses exceed 2%. If the loss threshold is exceeded, however, the full loss is accrued, not just the amount in excess of the 2% threshold.

Discussion: Performance in the PGP demonstration is measured by comparing a participating PGP's expenditure growth rate to a target or expected growth rate. Both the PGP and the target expenditure growth rates are subject to random fluctuations. Even if the PGP makes no changes in behavior in response to the demonstration, its performance (expenditure growth rate) may vary above or below the target growth rate due to random fluctuations. The goal of the PGP demonstration is to reward efficiency, not random fluctuations. Small differences between the PGP and target growth rates could well be due to chance; large differences are less likely to be due to chance. Two ways to increase the odds of paying a bonus rewarding efficiency are to: i) require a sufficiently large number of beneficiaries assigned to a participating PGP and its comparison group; and ii) pay a bonus only when the difference between the PGP and target growth rates exceeds a threshold. The former improves the accuracy of per capita cost estimation, increasing the chance that the bonuses paid are deserved. The latter avoids paying small bonuses that may be due to chance.

Minimum Number of Assigned Beneficiaries. We first discuss the number of assigned beneficiaries for participating PGPs and their comparison groups, then bonus thresholds. Table 7-1 is based on the simulation model developed in Pope and Chromy (1997). The left-hand-side columns show the assumed true percentage changes in the market area (comparison population) and participating site per capita expenditures.²³ In

²³ The true percentages are the average changes than can be expected to occur based on PGP cost control efforts. The true percentages are hypothetical and unobserved. The observed percentage changes may differ from the “true” (average/expected) changes due to random variations in expenditures. The observed percentage changes are not shown in Table 7-1, but the probability distribution of observed changes determines the probability of paying a bonus as shown in the table. The goal is to pay a bonus based on a PGP's true cost control efforts, but the bonus must be based on the (partly random) observed expenditures changes. The randomness of the observed percentage changes means that a bonus may sometimes be paid when not deserved, or not paid when deserved.

Table 7-1

Effect of PGP and Market Area Sample Sizes on the Probability of Paying a Bonus

			Bonus Threshold (in percents)									
			2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	
			Coefficient of Variation of Per Capita Expenditures									
Market Site			1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	1.40	
			Sample Size (number of assigned beneficiaries)									
Site			10,000	15,000	20,000	30,000	50,000	15,000	15,000	15,000	10,000	
Market Site			10,000	15,000	20,000	30,000	50,000	30,000	50,000	1,000,000	50,000	
<u>Percent Change in Market Per Capita Expenditures</u>	<u>Percent Change in PGP Site Per Capita Expenditures</u>	<u>Medicare Savings Rate Difference</u>	Serial Correlation of Per Capita Expenditures									
			Market Site	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
			Site	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
			Probability of Paying a Bonus in Percents									
10.0%	10.0%	0.0%		21.9%	17.1%	13.6%	8.9%	4.1%	13.6%	11.9%	9.1%	15.8%
8.0	8.0	0.0		21.5	16.6	13.2	8.5	3.9	13.2	11.5	8.7	15.4
6.0	6.0	0.0		21.0	16.2	12.7	8.2	3.6	12.7	11.1	8.3	14.9
4.0	4.0	0.0		20.6	15.7	12.3	7.8	3.3	12.3	10.6	7.9	14.5
6.0	10.0	-4.0		0.9	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.1
6.0	8.0	-2.0		5.5	2.5	1.2	0.3	0.0	1.2	0.7	0.3	1.9
6.0	6.0	0.0		21.0	16.2	12.7	8.2	3.6	12.7	11.1	8.3	14.9
6.0	4.0	2.0		50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
6.0	2.0	4.0		79.4	84.3	87.7	92.2	96.7	87.6	89.1	91.7	85.2
6.0	0.0	6.0		95.1	97.9	99.0	99.8	100.0	99.0	99.3	99.7	98.2

SOURCE: Health Economic Research, Inc.

the top four rows, the true difference in the site and market expenditure growth rates is zero, that is, the site is not controlling costs any better than its market area. In this case, a bonus payment is not justified. The bottom set of rows contains both scenarios in which the site is performing more poorly than the market (a negative true difference, market minus site, no bonus is appropriate) and scenarios in which the site is performing better than the market (a positive true difference, market minus site, a bonus payment is appropriate). The top of the table shows the assumed level of various simulation parameters, including the bonus threshold, the coefficient of variation of casemix-adjusted per capita Medicare expenditures, the market and site sample sizes (numbers of assigned beneficiaries), and the year to year correlation in expenditures among individual beneficiaries (the serial correlation). The table entries are the percentage probabilities of paying a bonus given the values of the parameters. The paid bonus is based on the observed site and market expenditure growth rates. On average, the observed rates will equal the true rates. But observed rates may differ from true rates due to random factors, and thus bonuses may be paid when not deserved and not paid when deserved. To discuss the number of assigned beneficiaries we assume a bonus threshold of 2% in Table 7-1, which we recommend below based on a separate analysis.

Given statistical fluctuations, it is impossible to expect to never pay an undeserved bonus in the PGP demonstration. The probability of paying an undeserved bonus can be minimized by only allowing extremely large PGPs to participate, but this would severely restrict the scope of the demonstration. In deciding on a minimum number of required PGP beneficiaries, it is useful to think about the highest probability

of paying an undeserved bonus that would be acceptable.²⁴ A reasonable position might be that paying an undeserved bonus to as many as 20% (one-fifth) of participating PGPs would be unacceptable, but that paying an undeserved bonus to 10% (one-tenth) would be acceptable. Bracketed by 10% and 20%, 15% is a reasonable acceptable maximum probability of paying an undeserved bonus. This would imply that in a demonstration with a large number of participating sites, about 15%, or one in 7, would be paid an undeserved bonus.

First consider the situation when PGP and comparison group number of beneficiaries (sample sizes) are equal. Table 7-1 shows that a sample size of 15,000 to 20,000 beneficiaries results in a 15% probability of paying an undeserved bonus (refer to the top panel of "probability of paying a bonus" in Table 7-1, when the difference between site and market area growth rates are zero).²⁵ However, it will often be the case that market area comparison group sample sizes are much larger than participating PGP sample sizes. For example, the far right column of Table 7-1 shows that a PGP sample size of 10,000 combined with a market area sample size of 50,000 also results in about a 15% probability of paying an undeserved bonus. This implies that the necessary PGP sample size can fall to as low as 10,000 or so with a large market area comparison group. In sum, HER recommends that CMS establish an assigned beneficiary minimum of 15,000 to 20,000 beneficiaries both for PGPs participating in the demonstration and their

²⁴ Symmetrically, one could think about the highest acceptable probability of not paying a bonus when it is deserved—Table 7-1 shows that the probabilities are roughly the same.

²⁵ Table 7-1 also shows that with a sample size of 15,000 to 20,000 the probability of paying a 4% bonus when it is deserved is about 84% (refer to the bottom panel of "probability paying a bonus" in Table 7-1 for a 4% difference between market and site expenditure growth rates), or the probability of not paying a bonus in this situation is 100% - 84% = 16%.

market area comparison groups. The PGP minimum could be lowered to as few as 10,000 assigned beneficiaries if 50,000 or more beneficiaries comprised the corresponding market area comparison group.

Bonus Threshold. As the bonus threshold is higher, the probability of not paying a bonus when it is not deserved increases (which is good), but the probability of not paying a bonus when it is deserved also rises (which is bad). Establishing the appropriate bonus threshold, therefore, involves balancing these two competing concerns. Table 7-2 simulates the probability of paying a bonus with two alternative bonus thresholds—2% and 4%—using the same simulation model from Pope and Chromy (1997) employed above to analyze the minimum number of required PGP and comparison group beneficiaries. For this simulation we specify three sample sizes for PGP and market area: 15,000 (small), 30,000 (medium), and 50,000 (large).

With a small PGP and market area (15,000 assigned beneficiaries), we see that the probability of paying an undeserved bonus is about 16% with a 2% threshold, but only about 2% with a 4% threshold (refer to top panel of "probability of paying a bonus" in Table 7-2 where site and market growth rates are equal). Clearly, the 4% threshold is more effective in preventing undeserved bonuses. However, the cost of this protection is that the probability of paying a deserved bonus is much lower with a 4% threshold (refer to the bottom panel of Table 7-2). When the deserved bonus is 2% (site growth rate less than market area growth rate by 2%), the probability of the PGP receiving a bonus with a 4% threshold is only 16% versus a 50% probability with a 2% threshold. With a 4% deserved bonus, the probability of paying it to the PGP are 50% (4% threshold) and 84%

Table7-2

Effect of Alternative Bonus Thresholds on Probability of Paying a Bonus

			Bonus Threshold (in percents)					
			2.00	4.00	2.00	4.00	2.00	4.00
			Coefficient of Variation of Variation of Per Capita Expenditures					
			1.40	1.40	1.40	1.40	1.40	1.40
			1.40	1.40	1.40	1.40	1.40	1.40
			Sample Size (number of assigned beneficiaries)					
			15,000	15,000	30,000	30,000	50,000	50,000
			15,000	15,000	30,000	30,000	50,000	50,000
<u>Percent Change in Market Per Capita Expenditures</u>	<u>Percent Change in PGP Site Per Capita Expenditures</u>	<u>Medicare Savings Rate Difference</u>	Serial Correlation of Per Capita Expenditures					
			0.3	0.3	0.3	0.3	0.3	0.3
			0.3	0.3	0.3	0.3	0.3	0.3
			Probability of Paying a Bonus in Percents					
10.0%	10.0%	0.0%	17.1%	2.9%	8.9%	0.4%	4.1%	0.0%
8.0	8.0	0.0	16.6	2.6	8.5	0.3	3.9	0.0
6.0	6.0	0.0	16.2	2.4	8.2	0.3	3.6	0.0
4.0	4.0	0.0	15.7	2.2	7.8	0.2	3.3	0.0
6.0	10.0	-4.0	0.2	0.0	0.0	0.0	0.0	0.0
6.0	8.0	-2.0	2.5	0.2	0.3	0.0	0.0	0.0
6.0	6.0	0.0	16.2	2.4	8.2	0.3	3.6	0.0
6.0	4.0	2.0	50.0	16.0	50.0	8.0	50.0	3.5
6.0	2.0	4.0	84.3	50.0	92.2	50.0	96.7	50.0
6.0	0.0	6.0	97.9	84.5	99.8	92.4	100.0	96.8

SOURCE: Health Economic Research, Inc.

(2% threshold). Similar comparisons can be made in Table 7-2 with larger PGP and market area sample sizes. As the number of beneficiaries assigned to the site and to the market area grows, the probability of inappropriately paying bonuses falls. Hence, a lower threshold can be used with larger site and market sample sizes to maximize the probability of paying a bonus when it is deserved. Given likely ranges of PGP and market area sample sizes for this demonstration, based on the results in Table 7-2, a bonus threshold of 2% seems reasonable.

A Note on the Definition of the Bonus Threshold. In this section, we have been discussing the bonus threshold as being defined in terms of the difference in site and market growth rates. An alternative definition of bonus threshold is in terms of the Medicare savings rate, defined as Medicare savings (target minus site expenditures) divided by target expenditures. These definitions are conceptually the same, but differ slightly in practice. The reason for the difference is that the growth rates use base year expenditures as the denominator of the percentage change and the Medicare savings rate uses performance year expenditures as the denominator of the percentage change.

An example shows the difference. Suppose base year per capita expenditures are \$5,000, the site expenditure growth rate is 8% and the market area expenditure growth rate is 10%. Then the difference in site and market expenditure growth rates is 2%. Target expenditures in the performance year are \$5,500 and site expenditures are \$5,400. So the Medicare savings rate is $(\$5,500 - \$5,400)/\$5,500 = 1.8\%$, which is less than the difference in growth rates of 2%. Note that if base year expenditures, \$5,000, were used as the denominator in the computation of the Medicare savings rate, that the Medicare savings rate would be 2%, the same as the difference in the growth rates. Note also that

the Medicare savings rate will always be lower than the difference in the growth rates as long as expenditures are rising (because performance year expenditures will exceed base year expenditures). So requiring the Medicare savings rate to exceed a 2% threshold is slightly tougher than requiring the difference in the growth rates to exceed 2%. As the expenditure growth rates increase, the Medicare savings rate will be increasingly smaller than the difference in expenditure growth rates. But for likely expenditure growth rates, the difference will not be large (generally 0.2% or less).

HER recommends that the bonus threshold be defined in terms of the Medicare savings rate, not the difference in site and market area expenditure growth rates. The reason for this recommendation is that bonus computations are carried out in terms of Medicare savings, not differences in growth rates. Also, it seems intuitive to think of the rate of savings from the demonstration as the Medicare savings divided by target expenditures, which are the amount expenditures are predicted to have been in the absence of the demonstration. Therefore, in the Executive Summary, we state the bonus threshold recommendation in terms of Medicare savings, not the difference in market and site expenditure growth rates.

Recommendation 8: Rebasing Expenditures

- PGP expenditures will not be rebased during the three year performance period of the demonstration.

Background: Rebasing means changing the base year. If the PGP demonstration is rebased annually, then each (performance) year becomes a new base year, and the performance of the participating PGP is judged independently on an annual basis. If the

demonstration is never rebased, the performance of the PGP is judged cumulatively over time. The rebasing issue can be equivalently interpreted in terms of over how long a period performance should be allowed to accumulate before the base year is "reset".

Table 8-1 shows an example of the effect of annual rebasing versus not rebasing during the course of the PGP demonstration. The demonstration has a base year (BY) and three successive performance years (PY1, PY2, PY3). The site's base year per capita expenditures are \$6,500, which are assumed to grow at 3% per year, compared to 8% for the comparison market area. The 8% market area growth rate applied to the site's base expenditures defines the target. The difference between the site's actual per capita expenditures and the target is the Medicare Savings. A sharing rate of 80% is assumed, so the PGP's bonus pool is 80% of the Medicare Savings. The remainder accrues as savings to the Medicare program.

Without rebasing, the 8% market growth rate is applied cumulatively each year to the site's base year expenditures. That is, the first year target is $(1.08) * (\$6,500)$, the second year target is $(1.08) * (1.08) * (\$6,500)$, and the third year target is $(1.08) * (1.08) * (1.08) * (\$6,500)$. Under annual rebasing, the base year is updated each year. The base year for PY1 is the PGP's base year expenditures, but the base year for PY2 is the PGP's PY1 expenditures, and the base year for PY3 is its PY2 expenditures. Because the site's expenditures are growing less rapidly than its market area, target expenditures grow less rapidly under annual rebasing. Target expenditures for PY1 are

Table 8-1

Effect of Annual Versus No Rebasing on Per Capita PGP Bonus Pool and Medicare Program Savings

Year	Expenditure Growth Rate <u>PGP Site</u>	Comparison <u>Group</u>	<u>No Rebasing</u>					<u>Annual Rebasing</u>				
			<u>Expenditures</u> <u>PGP Site</u>	<u>Target</u>	<u>Medicare</u> <u>Savings</u>	<u>PGP</u> <u>Bonus</u> <u>Pool</u>	<u>Medicare</u> <u>Program</u> <u>Savings</u>	<u>Expenditures</u> <u>PGP Site</u>	<u>Target</u>	<u>Medicare</u> <u>Savings</u>	<u>PGP</u> <u>Bonus</u> <u>Pool</u>	<u>Medicare</u> <u>Program</u> <u>Savings</u>
BY	--	--	\$6,500	--	--	--	--	\$6,500	--	--	--	--
PY1	3%	8%	6,695	\$7,020	\$325	\$260	\$65	6,695	\$7,020	\$325	\$260	\$65
PY2	3	8	6,896	7,582	686	549	137	6,896	7,231	335	268	418
PY3	3	8	7,103	8,188	1,085	868	217	7,103	7,448	345	276	810

NOTES:

BY = Base Year. PY = Performance Year.

Assumes a sharing rate = 0.8

Medicare Program Savings under annual rebasing are calculated by subtracting the PGP bonus pool from Medicare Savings without rebasing.

SOURCE: Health Economics Research, Inc.

the same, but the target for PY2 is $(1.08)*(\$6,695) = \$7,231$ which is less than the target of \$7,582 with no rebasing.

What are the implications of rebasing versus not rebasing? As shown in Table 8-1, the PGP's bonus pool grows much more rapidly without rebasing. In PY3, the bonus pool is \$868 per capita without rebasing compared to only \$276 with rebasing. With annual rebasing, the PGP's bonus pool remains virtually constant over time, because performance is judged year by year. Without rebasing, the PGP is cumulatively rewarded for past performance in each performance year after the first one. Conversely, Medicare program savings are greater with annual rebasing. They grow from \$65 to \$810 under annual rebasing²⁶, but only from \$65 to \$217 without rebasing. Note that the sum of the PGP bonus pool and Medicare program savings is the same under either annual rebasing or no rebasing. Rebasing allocates more of the total savings to the Medicare program, while not rebasing allocates more of it to the participating PGP.

Discussion: If the participating site holds its rate of expenditure growth below the target in the first performance year of the demonstration, its per capita expenditures will remain below its target indefinitely even if it just matches comparison group expenditure growth rates for all other performance years. Without rebasing, a PGP could indefinitely earn a bonus based on holding its cost growth below target in one year. One could argue that if continuing management of care (e.g., disease management programs) is required to continue to keep the level of expenditures lower (e.g., through continuing

²⁶ Under annual rebasing, Medicare program savings are calculated by subtracting the PGP bonus pool from Medicare Savings with no rebasing. Medicare Savings without rebasing measures the actual annual savings from the lower rate of simulated cost growth of the participating PGP.

to keep the hospital admission rate at a lower level), then a permanent lower level of expenditures should be rewarded, even if the rate of growth of expenditures reverts to the comparison amount.

On the other hand, one could argue that a participating PGP should not be rewarded indefinitely for having made a one-time lowering of its costs, especially if it was starting from an inflated, inefficient level. This perspective would argue that additional bonuses should be paid only if a PGP makes further reductions in cost every year. Under this perspective, cost-reducing efforts by the PGP are seen as one-time expenditures that only need to be rewarded once, not indefinitely. More frequent rebasing will capture more of the total cost savings for Medicare.²⁷

A compromise position would be to rebase periodically, but not every year. Costs might be rebased every 3 years, for example. Bonuses would be paid as in the no rebasing scenario described above for 3 years, but then the base year and expenditures would be reset. This policy would give PGPs a greater incentive to lower costs than under annual rebasing, but would capture more of the aggregate cost savings for Medicare over time.

Another aspect to the rebasing issue is allowing cost savings to compound to exceed the bonus threshold. If costs are rebased annually, then it is less likely that a PGP will be able to exceed its threshold and ever earn a bonus. For example, suppose the bonus threshold is 2% and the participating PGP holds its expenditure growth rate 1

²⁷ The sharing ratio means that even with no rebasing, Medicare will share in some of the savings generated by a PGP that creates Medicare savings and earns bonuses. However, Medicare will capture even more of the total savings if it rebases frequently.

percentage point below the target expenditure growth rate each year. With annual rebasing, this PGP will never earn a bonus. But with less frequent rebasing, its cost savings will eventually compound, allowing it to earn a bonus. For example, after 3 years, its expenditures will be 3% below its target, which exceeds the 2% threshold and will earn it a bonus.

Over the relatively short period of the demonstration (3 years), HER recommends that rebasing should not occur. Given the potential random element in annual bonus calculations, it is better to let Medicare savings and losses accumulate over this short period so that losses can offset gains. Further, allowing cost savings to compound over this period will allow annual savings to eventually exceed the bonus threshold. Exceeding the threshold is more difficult with annual rebasing.

Recommendation 9: Preliminary Expenditure Targets and Interim Performance Reporting

- Preliminary expenditure targets will not be announced in the PGP demonstration. But CMS will provide participating PGPs with interim reports on hospital utilization so that they may gauge their performance on a timely basis.

Background: In the proposed PGP demonstration, Medicare savings are calculated by comparing a PGP's actual performance year per capita expenditures to target per capita expenditures. Target expenditures are computed by applying per capita expenditure growth in a comparison population to base year participating PGP adjusted per capita expenditures. Neither actual nor target expenditures are known until well after the end of the performance year. The claims data necessary to compute the actual and target expenditures will not become available in relatively complete form until at least 6

months after the end of the performance year. Allowing time for obtaining and processing data, the actual expenditures, final target expenditures, and the bonus will not be known until at least 8 to 12 months after the end of the performance year.

Given these lags in measuring performance, preliminary expenditure targets or performance reporting could be useful to the participating PGP so that it can judge and adjust its performance in a timely fashion. For the reasons discussed next, we do not believe preliminary expenditure targets are feasible or desirable. But we recommend that 6 month interim reports on utilization be provided to participating PGPs.

Discussion: A preliminary target has several disadvantages. First, it may not be available before the end of the performance year, and thus could not be a standard to which the PGP manages. The simplest way to compute the preliminary target would be to trend forward base year expenditures. But assuming the base year immediately precedes the performance year, the claims necessary to compute base year expenditures would not be available until 6 months into the performance year. With data request and processing lags, the preliminary target would not be available until near the end of the performance year. The preliminary target could be based on data from the year before the base year, but this would significantly increase administrative burden by requiring processing another year of data; also a preliminary target based on older data would be less accurate. Moreover, a preliminary target will not allow an earlier estimate of a PGP's bonus because a PGP's own expenditure performance won't be known until 8-12 months after the end of the performance year.

Second, the preliminary and final targets are often likely to differ significantly. Expenditure growth rates are notoriously hard to predict, especially for local markets.

CMS sometimes will be put in the awkward position of making major revisions in its target and therefore in the participating PGP's bonus. For example, the preliminary target might show the PGP earning a substantial bonus whereas the final target might indicate no bonus at all. Large differences between preliminary and final targets will undermine participating PGPs' confidence in the accuracy of CMS payment calculations and in the fairness of the demonstration.

Third, calculating the preliminary target could impose a substantial administrative burden. If either a) an earlier year than the base year is used to compute the preliminary target, or b) the PGP demonstration contractor has to compute the market growth rate projections based on multi-year trends in local expenditure growth, additional years of data would have to be processed. This would be time consuming and expensive.

Although we recommend that CMS does not announce preliminary expenditure targets, interim reporting on hospital utilization is feasible. The goal of interim reporting is to give participating sites an indication of how well they are managing care and performing in controlling utilization. We recommend that interim reporting be done for the first six months of a performance year. Because timeliness of interim reports is the goal, we would allow only a 3 month claims lag for interim reporting. We believe that inpatient claims should be mostly complete after 3 months. Allowing some time for data processing and report writing, the interim reports should be available by the end of the performance year. For example, the interim report for the first six months of the first performance year will be available at the end of the first performance year. Although this will be too late for the PGP to modify its behavior in the first performance year, it will be able to modify its performance in the second performance year. We also considered

quarterly interim reporting, but rejected it as too administratively burdensome, and too short a time period to serve as a useful or statistically reliable performance indicator.

A challenge in interim reporting is how to identify beneficiaries assigned to the participating PGP. For the final performance year computations, involved beneficiaries are identified through physician utilization over the course of the entire performance year. But this universe of involved beneficiaries cannot be known before the end of the performance year. The two possibilities for interim reporting are to calculate performance measures for beneficiaries assigned to the PGP in the base (or prior) year, or to assign beneficiaries from the first six months of performance year data. Both of these options have drawbacks, but the former is more feasible and preferable. Base year assignees represent a biased population in the performance year because they exclude new enrollees assigned in the performance year to replace beneficiaries who died in the base year. Comparisons of utilization of the base year population for the base and performance years will need to be adjusted for mortality and new enrollees. But base year assignees will already be generated for bonus computations; using this population for interim reporting will not create any additional administrative burden. Conversely, assigning beneficiaries from the first six months of performance year data will create substantial additional administrative burden by requiring an entire additional data processing cycle. Sample sizes will also be smaller with only beneficiaries assigned based on 6 months of data, and the sample will be biased because of the use of only a half-year of data to assign beneficiaries.

A template for interim performance reporting for a hypothetical participating PGP is given in Table 9-1. Hospital inpatient expenditures are the largest single expenditure category (accounting for roughly half of total expenditures), and are expected to be a primary focus of site cost control efforts. Hospital admissions are relatively easily measured and counted. We recommend that interim utilization reporting to sites focus on rates of hospital admissions (discharges). In particular, we recommend that rates of admissions for ambulatory care sensitive conditions (discussed in Part V of this report) be reported, both as an overall index and for specific conditions such as congestive heart failure. These are hospitalizations that are thought often to be avoidable, and thus may be a primary focus of cost control efforts. Moreover, they are one of the quality of care measures that we propose (see Part V).

Interim utilization will be compared for the PGP's assigned base year population between the first six months of performance and base years. We will also compute hospital admission rates for the PGP's base year market area comparison group in the first six months of the base and performance years. These can be compared to the change in the PGP's own hospital utilization for an indication of whether the participating PGP is controlling utilization better than its comparison group. Finally, we will provide participating PGPs with per capita expenditures (a weighted average of county amounts) for the PGP's market area.

Table 9-1

Template for Interim Utilization Reporting (for a Physician Group Practice)

	Participating PGP			Market Area		
	<u>First Six Months of:</u>			<u>First Six Months of:</u>		
	<u>Base Year</u>	<u>Performance Year</u>	<u>% Change</u>	<u>Base Year</u>	<u>Performance Year</u>	<u>% Change</u>
Total admission rate, all conditions						
Ambulatory Care Sensitive Condition						
Admission Rates per 1,000 Beneficiaries						
Total index						
Congestive heart failure						
Chronic obstructive pulmonary disease						
Pneumonia						

SOURCE: Health Economics Research, Inc.

Part IV

Expenditure Definition and Adjustments

IV. Expenditure Definition and Adjustments

Recommendation 10: Expenditure Definition

- Use all Medicare Part A and Part B expenditures to calculate per capita expenditures for the demonstration. Since the primary goal of the PGP demonstration is to reduce the growth rate of Medicare fee-for-service expenditures, setting a comprehensive target gives the PGP more flexibility to focus on the largest sources of inefficiency. It thus encourages PGPs to take a comprehensive view when managing care and developing ways to better coordinate care.

Background: BIPA 2000 requires that the PGP demonstration include "a base expenditure amount, equal to the average total payments under parts A and B for patients served by the health care group on a fee-for-service basis in a base period determined by the Secretary". Although BIPA 2000 is explicit about including "total payments under parts A and B", there may be some flexibility to exclude certain minor part A and/or part B expenditures if they would significantly raise administrative burden or delay payment settlements.

Discussion: All Medicare expenditures is the most comprehensive basis for computing the PGP base expenditure amounts, and is most consistent with the BIPA 2000 requirement. The components of Medicare expenditures (claims) are as follows:

- Hospital Inpatient;
- Skilled Nursing Facility (SNF);
- Hospital Outpatient;
- Physician/Supplier Part B;
- Hospice;
- Home Health Agency (HHA);
- Durable Medical Equipment (DME).

Each component is available roughly six months into the following year. Thus there is no argument for excluding any component of Medicare claims based on data lags that would delay the calculation of demonstration bonuses. Further, given that no adjustments are made for changes in Medicare payment policies (Recommendation 14), it is not significantly more administratively burdensome to compute expenditures that incorporate all components of Medicare claims than it is to compute expenditures that incorporate only a subset of Medicare claims (e.g., hospital inpatient and physician/supplier Part B). Because of substitution among expenditure types, basing the expenditure amounts on only part of expenditures or utilization could lead to inappropriate and unfair bonus payments. For example, suppose PGP A substituted SNF services for hospital inpatient services whereas PGP B did not. If SNF expenditures were excluded from the target as a minor expenditure category, PGP A might earn a bonus when it has not reduced overall expenditures. Moreover, PGP A might unfairly earn a larger bonus than PGP B. A more narrowly defined expenditure target will also be less stable statistically than a target based on all expenditures.

In a 1998 HER report for CMS²⁸ entitled *Simulations of Selected GVPS Design Parameters Using Historical Data*, data for seven PGPs are used to compute historical growth rates (from 1993 to 1994) for simulated PGP sites and state market areas by type of service. Results from this report are shown in Table 10-1. As shown in Table 10-1,

²⁸ See Adamache, Liu, and Pope, *Simulations of Selected GVPS Design Parameters Using Historical Data*, Technical Report, CMS Contract No. 500-95-0048, March 11, 1998.

Table10-1

Per Capita Medicare Payments¹ and Growth Rates, 1993 and 1994, by Medical Service Category

Potential PGP Sites	Mean Payment, by Site				Mean Payment, by State ²				
	1993	1994	Growth Rate	Percentage of Total Growth	1993	1994	Growth Rate	Percentage of Total Growth	
Combined Sites	Parts A + B	\$7,253	\$7,706	6.2 %	\$4,614	\$4,990	8.1 %		
	PPS Hospital	3,755	3,828	1.9	16.1 %	2,185	2,233	2.2	12.8 %
	PPS Exempt Hospital	302	356	17.9	12.0	192	213	11.1	5.7
	SNF	147	210	43.3	14.0	147	200	35.7	14.0
	Home Health	350	419	19.8	15.3	343	442	28.9	26.4
	Hospice	42	50	18.8	1.7	39	51	29.2	3.1
	Total Part A	4,596	4,863	5.8	59.1	2,907	3,139	8.0	61.9
	Physician	1,540	1,673	8.6	29.2	970	1,064	9.7	25.0
	Supplies & Other	227	205	-9.6	-4.8	200	183	-8.8	-4.7
	DME	133	155	16.2	4.8	104	136	31.5	8.7
	Outpatient	757	810	6.9	11.6	433	468	8.0	9.2
	Total Part B	2,658	2,842	7.0	40.8	1,707	1,850	8.4	38.1
PGP A	Parts A + B	9,333	10,719	14.9		5,555	5,896	6.1	
	PPS Hospital	4,571	5,209	13.9	46.0	2,578	2,522	-2.2	-16.3
	PPS Exempt Hospital	1,171	1,257	7.4	6.2	386	427	10.8	12.2
	SNF	394	548	39.2	11.1	196	286	45.7	26.3
	Home Health	718	819	14.0	7.3	522	636	22.0	33.6
	Hospice	15	17	9.0	0.1	37	43	17.1	1.9
	Total Part A	6,869	7,849	14.3	70.7	3,719	3,915	5.3	57.6
	Physician	1,384	1,588	14.7	14.7	946	1,027	8.6	23.8
	Supplies & Other	413	441	7.0	2.1	279	279	-0.2	-0.1
	DME	122	172	41.4	3.6	85	106	24.8	6.1
	Outpatient	545	669	22.7	8.9	526	569	8.2	12.6
	Total Part B	2,463	2,870	16.5	29.3	1,836	1,981	7.9	42.4
PGP B	Parts A + B	7,033	8,092	15.1		5,555	5,896	6.1	
	PPS Hospital	3,453	3,715	7.6	24.7	2,578	2,522	-2.2	-16.3
	PPS Exempt Hospital	502	710	41.5	19.7	386	427	10.8	12.2
	SNF	142	305	114.6	15.4	196	286	45.7	26.3
	Home Health	460	675	46.5	20.2	522	636	22.0	33.6
	Hospice	32	44	39.4	1.2	37	43	17.1	1.9
	Total Part A	4,589	5,448	18.7	81.2	3,719	3,915	5.3	57.6
	Physician	1,252	1,377	10.0	11.8	946	1,027	8.6	23.8
	Supplies & Other	231	269	16.4	3.6	279	279	-0.2	-0.1
	DME	76	101	32.5	2.3	85	106	24.8	6.1
	Outpatient	885	897	1.3	1.1	526	569	8.2	12.6
	Total Part B	2,444	2,643	8.2	18.8	1,836	1,981	7.9	42.4

Table 10-1 (continued)

Per Capita Medicare Payments¹ and Growth Rates, 1993 and 1994, by Medical Service Category

Potential PGP Sites	Mean Payment, by Site				Mean Payment, by State ²				
	1993	1994	Growth Rate	Percentage of Total Growth	1993	1994	Growth Rate	Percentage of Total Growth	
PGP C	Parts A + B	5,502	5,773	4.9		4,947	5,268	6.5	
	PPS Hospital	2,643	2,682	1.5	14.5	2,387	2,469	3.5	25.8
	PPS Exempt Hospital	273	288	5.7	5.7	256	271	5.7	4.6
	SNF	135	177	31.6	15.7	154	201	30.2	14.5
	Home Health	385	431	11.8	16.8	284	347	22.1	19.6
	Hospice	27	31	13.2	1.3	18	26	41.5	2.3
	Total Part A	3,463	3,609	4.2	54.0	3,100	3,315	6.9	66.9
	Physician	1,148	1,227	6.9	29.3	1,087	1,171	7.7	26.1
	Supplies & Other	172	144	-16.1	-10.2	215	191	-11.3	-7.6
	DME	103	128	24.4	9.3	96	130	35.4	10.6
	Outpatient	617	665	7.7	17.7	449	462	2.9	4.1
	Total Part B	2,040	2,164	6.1	46.0	1,847	1,953	5.8	33.1
PGP D	Parts A + B	10,013	10,112	1.0		4,361	4,768	9.3	
	PPS Hospital	5,760	5,513	-4.3	-250.0	2,215	2,197	-0.8	-4.5
	PPS Exempt Hospital	258	310	20.2	52.6	162	170	5.2	2.1
	SNF	137	216	58.3	80.3	143	200	40.1	14.1
	Home Health	412	547	32.7	135.9	207	280	35.0	17.8
	Hospice	64	74	15.1	9.7	44	59	33.4	3.6
	Total Part A	6,631	6,660	0.4	28.5	2,772	2,907	4.9	33.1
	Physician	1,966	2,026	3.1	60.5	901	985	9.3	20.5
	Supplies & Other	245	195	-20.4	-50.5	166	148	-10.9	-4.4
	DME	161	212	31.6	51.2	90	123	37.5	8.2
	Outpatient	1,010	1,020	1.0	10.3	433	458	5.8	6.2
	Total Part B	3,382	3,453	2.1	71.5	1,589	1,861	17.1	66.9
PGP E	Parts A + B	8,095	8,399	3.7		4,589	4,970	8.3	
	PPS Hospital	3,980	4,063	2.1	27.3	2,212	2,278	3.0	17.4
	PPS Exempt Hospital	368	394	6.8	8.3	180	196	8.7	4.1
	SNF	162	202	24.3	13.0	89	130	45.6	10.7
	Home Health	437	437	0.1	0.1	257	337	31.5	21.2
	Hospice	60	68	13.1	2.6	50	70	41.3	5.4
	Total Part A	5,007	5,163	3.1	51.3	2,788	3,012	8.0	58.8

Table 10-1 (continued)

Per Capita Medicare Payments¹ and Growth Rates, 1993 and 1994, by Medical Service Category

Potential PGP Sites	Mean Payment, by Site				Mean Payment, by State ²			
	1993	1994	Growth Rate	Percentage of Total Growth	1993	1994	Growth Rate	Percentage of Total Growth
PGP E (Continued)								
Physician	1,480	1,611	8.8	43.1	1,022	1,135	11.0	29.5
Supplies & Other	261	221	-15.4	-13.2	173	162	-6.1	-2.8
DME	227	193	-15.0	-11.2	129	136	5.8	2.0
Outpatient	1,119	1,210	8.1	30.0	478	525	9.9	12.5
Total Part B	3,088	3,236	4.8	48.7	1,802	1,958	8.7	41.2
PGP F								
Parts A + B	6,549	7,062	7.8		3,513	3,602	2.5	
PPS Hospital	3,662	3,787	3.4	24.4	1,910	1,832	-4.1	-88.2
PPS Exempt Hospital	174	208	19.8	6.7	110	114	3.6	4.4
SNF	131	181	37.7	9.6	196	213	8.6	19.1
Home Health	151	183	21.3	6.2	114	131	14.8	19.2
Hospice	30	40	33.9	2.0	19	45	135.6	29.1
Total Part A	4,148	4,399	6.1	49.0	2,349	2,334	-0.6	-16.3
Physician	1,644	1,811	10.1	32.4	645	714	10.7	78.7
Supplies & Other	169	151	-10.5	-3.5	95	87	-8.2	-8.8
DME	99	137	37.7	7.3	53	67	26.0	15.6
Outpatient	489	564	15.5	14.7	371	399	7.3	30.8
Total Part B	2,401	2,663	10.9	51.0	1,165	1,267	8.8	116.3
PGP G								
Parts A + B	6,727	6,885	2.4		5,092	5,664	11.2	
PPS Hospital	2,853	2,781	-2.5	-45.3	1,996	2,102	5.3	18.5
PPS Exempt Hospital	123	202	64.5	49.9	174	193	10.9	3.3
SNF	127	150	18.4	14.7	172	252	46.5	14.0
Home Health	317	373	17.8	35.5	504	616	22.3	19.7
Hospice	63	61	-4.4	-1.8	66	83	24.5	2.8
Total Part A	3,482	3,566	2.4	53.1	2,912	3,246	11.4	58.3
Physician	2,180	2,239	2.7	37.5	1,320	1,484	12.4	28.7
Supplies & Other	414	345	-16.7	-43.6	362	320	-11.6	-7.3
DME	121	161	33.2	25.3	110	189	72.4	13.9
Outpatient	530	574	8.3	27.7	389	425	9.4	6.4
Total Part B	3,244	3,319	2.3	46.9	2,180	2,418	10.9	41.7

¹ Medical Payments are annualized and weighted by the number of months on Medicare.

² The values for "combined sites" are national values.

there are large variations in expenditure growth rates by type of service and across sites. For combined sites, growth in PPS hospital expenditures is 1.9%, whereas growth in SNF expenditures is 43.3%. The large growth in post-acute care expenditures may have been substituting for hospital inpatient services. These data show the danger in excluding "minor" expenditure categories that may grow rapidly while substituting for major included categories.

Since the primary goal of the PGP demonstration is to reduce the growth rate of Medicare fee-for-service expenditures, setting a comprehensive target gives the PGP more flexibility to focus on the largest sources of inefficiency. It thus encourages PGPs to take a comprehensive view when managing care and developing ways to better coordinate care.

Recommendation 11: Medicare Pass-Throughs and PPS Add-Ons

- PPS pass-throughs and add-ons will be included when calculating per capita expenditures for the demonstration.

Background. Pass-throughs to PPS hospitals include payments for direct graduate medical education (DGME), bad debt, costs of acquiring organs for transplants, and pass-through capital costs. PPS add-ons include per discharge payments to qualifying hospitals serving a disproportionate share of poor patients (DSH) and indirect medical education (IME) payments to hospitals with residency programs.

In the design for the older GVPS demonstration it was recommended that neither pass-throughs nor add-ons be included in calculating the overall size of the bonus pool, per capita expenditures for both sites and comparison areas, and the comparison growth

rate (Tompkins *et al.*, 1996a). However, it was recommended that the sharing rate, one of the factors used to divide the bonus pool between the PGP and Medicare, include an adjustment for the PPS add-ons:

$$\text{Sharing Rate} = 0.75 + [(\text{IME}_{j,p} + \text{DSH}_{j,p}) \div \text{MR}_{j,p}]$$

where *j* denotes the GVPS organization and *p* denotes the performance year, and $\text{MR}_{j,p}$ is the sum of all Medicare reimbursements. Participating groups without a teaching hospital component and not receiving DSH payments would have, for example, a minimum 0.75 share. The DSH and IME payments included in the sharing formula were only for those GVPS organizations that have affiliated hospitals that already receive DSH and IME payments. For demonstration participants with an affiliated hospital that had a combined twelve percent IME and DSH share of Medicare reimbursements, the sharing rate would be increased from 0.75 to 0.87. The total sharing rate should be capped at a value such as 0.95 so that Medicare would receive at least a significant share of the savings.

Discussion. The design for the older GVPS demonstration recommended excluding IME and DSH from per capita expenditures and comparison growth rate because “the focus of the demonstration is utilization management, not IME or DSH funding policies.” (Tompkins *et al.*, 1996a) IME and DSH payment adjustment rates have been, historically, higher than the necessary to compensate for the associated costs of IME and providing services to poor patients. Congress, presumably, provided hospitals this financial safety margin so that hospitals could continue to provide graduate medical education and to serve poor populations.

We conducted simulations of the compensation to teaching hospitals for forgone add-on payments using the old GVPS method (see Appendix E).²⁹ We conclude that under currently proposed demonstration rules as described in this report, the old GVPS method overcompensates teaching hospitals for their foregone IME and DSH payments. Hence, we recommend that pass-throughs and PPS add-ons be treated like any other expenditure. That is, include them when calculating per capita expenditures, the target growth rate, total (Medicare) savings, and the overall size of the bonus pool. The participating PGP with a teaching hospital can get up to 80 percent of the foregone add-ons and passthroughs back through the bonus payment (exact amount depending on quality performance). At least 20 percent of the foregone amounts would be forfeited to the government as a Medicare program savings. But this 20 percent of "lost" revenue could be covered out of the remaining bonus payment and cost savings from fewer admissions.

With this approach, PGPs could increase their bonus by shifting admissions from teaching to community hospitals. But they would forfeit the IME and (potentially) DSH payments by doing this. Whether it would be profitable for them to move admissions would depend on relative costs in the teaching and community hospitals. This approach is consistent with Medicare policy (including BBA treatment of teaching pass-throughs in M+C) of treating IME and DSH as variable costs and paying for them on a per discharge basis.

²⁹ Appendix E also discusses the issue of the timeliness of settlement for IME payments, and concludes that for the demonstration, it will be acceptable to use estimated IME payment amounts recorded on Medicare claims.

Recommendation 12: Cost Outliers

- For each beneficiary assigned to a PGP or comparison group, truncate (cap) annualized expenditures at \$100,000.

Background: The PGP demonstration is intended to give PGPs incentives to lower the volume and intensity of services provided to the beneficiaries assigned to them. PGPs that generate Medicare savings are paid a bonus, and those that do not are denied a bonus payment. Random variability of expenditure growth rates for PGP demonstration participants may reward PGPs even when they have not altered their behavior, and conversely, might lead to a lack of bonuses even when participants are reducing services per beneficiary.

Discussion: In a 1998 HER report for CMS³⁰ entitled *Simulations of Selected GVPS Design Parameters Using Historical Data*, 1993-1994 data for seven PGPs are used to determine the effect of eliminating cost outliers on the difference in growth rates between the PGPs and their comparison groups. Results from that report are presented in Table 12-1. As shown in Table 12-1, the exclusion of high cost outliers has a limited impact on the difference in the growth rates between PGPs and their state market areas. In particular, the exclusion of a few extremely high cost outliers (such as beneficiaries with expenditures of more than \$500,000 or \$1,000,000) has a very minor impact on the differences in growth rates.

Nevertheless, a small number of costly beneficiaries could significantly change a PGP's per capita expenditures. For example, consider a PGP with 20,000 assigned

³⁰ See Adamache, Liu, and Pope, *Simulations of Selected GVPS Design Parameters Using Historical Data*, Technical Report, CMS Contract No. 500-95-0048, March 11, 1998.

Table 12-1

Percent Change in Mean Medicare Payments, 1993-1994 (Excluding High Cost Users)

Potential PGP Sites	Percent Change in Mean Medicare Payments, by PGP Site, 1993-1994				Percent Change in Mean Medicare Payments, by State, 1993-1994*				Mean Difference in Percent Change Between State and PGP Site				
	<i>All Users</i>	Excluding Users with Expenditures More than or Equal to:			<i>All Users</i>	Excluding Users with Expenditures More than or Equal to:			<i>All Users</i>	Excluding Users with Expenditures More than or Equal to:			
		<i>\$100K</i>	<i>\$500K</i>	<i>\$1.0 M</i>		<i>\$100K</i>	<i>\$500K</i>	<i>\$1.0 M</i>		<i>\$100K</i>	<i>\$500K</i>	<i>\$1.0 M</i>	
Combined													
Parts A + B	6.2 %	5.8 %	6.2 %	6.2 %	8.1 %	7.8 %	8.1 %	8.1 %	1.9	2.0	1.9	1.9	
PGP A													
Parts A + B	14.9	13.4	14.9	14.9	6.1	7.4	6.5	6.2	-8.7	-6.0	-8.4	-8.6	
PGP B													
Parts A + B	15.1	12.2	14.8	15.0	6.1	7.4	6.5	6.2	-8.9	-4.8	-8.3	-8.8	
PGP C													
Parts A + B	4.9	4.0	5.0	4.9	6.5	6.0	6.3	6.5	1.6	2.0	1.3	1.6	
PGP D													
Parts A + B	1.0	2.1	0.9	1.0	5.9	6.1	5.9	5.9	5.0	4.0	5.1	5.0	
PGP E													
Parts A + B	3.7	2.9	3.8	3.7	8.3	8.4	8.3	8.3	4.6	5.5	4.5	4.5	
PGP F													
Parts A + B	7.8	7.5	7.9	7.8	2.5	3.0	2.7	2.4	-5.3	-4.5	-5.3	-5.4	
PGP G													
Parts A + B	2.4	2.5	2.3	2.4	11.2	9.8	11.2	11.2	8.9	7.3	8.9	8.9	

* The values in the rows for "combined sites" are national values.

SOURCE: KW Adamache, CF Liu, and GC Pope. "Simulations of Selected GVPS Design Parameters Using Historical Data", Technical Report, CMS Contract Number 500-95-0048, March 11, 1998.

beneficiaries and per capita expenditures of \$6,500. If 100 of those beneficiaries, or 0.5%, had annualized expenditures of \$100,000, per capita expenditures would rise by \$500, or 8% ($100 \times 100,000 / 20,000 = \500 , which is 8% of \$6,500). This could significantly affect growth rate and bonus calculations. Thus, HER recommends that annualized expenditures be capped \$100,000 per beneficiary for the PGP demonstration.

In an HER project for CMS³¹ entitled *Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment*, the distribution of 1997 annualized expenditures for fee-for-service (FFS) beneficiaries is computed using the 5% Standard Analytic File. The results are shown in Table 12-2. As shown in Table 12-2, more than 99% of beneficiaries had expenditures less than \$100,000, which is where cost outliers are capped for the PGP demonstration. Note that a primary objective of the PGP demonstration is to give PGPs an incentive to coordinate and manage the health care of the high cost patients assigned to them. Capping expenditures at a level significantly lower than \$100,000 runs the risk of jeopardizing this objective.

Recommendation 13: Health Status Casemix Adjustment

- The per capita expenditures of both the participating PGPs and their comparison groups will be adjusted for health status casemix using the concurrent DCG-HCC model.

Background: The PGP demonstration is intended to reward improvements in the efficiency of medical practice. It does this by creating a bonus pool based on the rate of

³¹ See Pope *et al.*, *Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment*, CMS Contract No. 500-95-048, December 21, 2000.

Table 12-2

Distribution of Annualized 1997 Medicare Expenditures¹

N	1,394,701
Mean ²	\$5,314
Standard Deviation ²	\$13,822
Coefficient of Variation ²	260%
Standard Error ²	\$12
<u>Percentile</u>	
Max ³	\$1,997,706
99	78,748
95	31,437
90	17,142
75	4,114
50 (Median)	844
25	189
10	4
5	0
1	0
Min	0
% non-users (zero payments)	9.7%

¹ For prospective modelling sample.

² Weighted by fraction of year alive.

³ The maximums shown is of annualized expenditures. The maximum of actual expenditures was \$566,302.

SOURCE: Gregory C. Pope, et al. "Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment", CMS Contract Number 500-95-048, December 21, 2000.

growth in Medicare spending per participating beneficiary compared to the growth in per capita spending in a comparison group of beneficiaries. But the rate of growth in per capita spending can be affected by changes in casemix, or the expected costliness, of the beneficiaries in a group. BIPA requires that the performance targets be adjusted for "risk".

Discussion: We recommend use of the concurrent Diagnostic Cost Groups, Hierarchical Condition Categories (DCG-HCC) model (Pope *et al.*, 2000) for casemix adjustment in the PGP demonstration. This model was developed by Health Economics Research, Inc., under contract to CMS, and is a member of the same family of models as the PIP-DCG model currently used for risk adjustment of capitation payments to Medicare+Choice plans. The DCG-HCC model uses demographic and eligibility information, and current year diagnoses on administrative claims to predict Medicare expenditures.

The concurrent DCG-HCC model differs from the PIP-DCG model used for M+C risk adjustment in two key respects. First, the PIP-DCG model utilizes only inpatient diagnoses while the DCG-HCC model uses inpatient, hospital outpatient, physician, and clinically-trained non-physician diagnoses. The only reason the PIP-DCG model is limited to inpatient data is that only inpatient encounters are collected from M+C plans. Since ambulatory as well as inpatient claims are available for Medicare fee-for-service enrollees, there is no reason to restrict risk adjustment to inpatient-only models in the fee-for-service context. Incorporating ambulatory claims, as the DCG-HCC model does, substantially improves the predictive accuracy of risk adjustment (Pope *et al.*, 2000).

Also, PGPs that reduce expenditures by avoiding unnecessary hospitalizations will not be penalized because beneficiary diagnoses will still be captured on ambulatory claims.

The second major difference is that we recommend a concurrent rather than prospective risk adjustment model for the PGP demonstration. The model used for M+C payment is "prospective", that is, it uses this year's diagnoses to predict next year's expenditures. In contrast, the "concurrent" model we recommend uses this year's diagnoses to predict this year's expenditures. The basic difference between prospective and concurrent models is that prospective models place more emphasis on chronic illnesses that have predictably higher future costs whereas concurrent models better explain acute illnesses that have higher costs this year. Because they capture acute illness much better than prospective models, concurrent models explain 40-50% of annual per capita expenditure variation among beneficiaries whereas prospective models explain only 10% (Ellis, Pope, *et al.*, 1996).

Prospective risk adjustment models are used for forecasting future expenditures (i.e., setting prospective capitation rates) based on current year beneficiary information. In contrast, the need in the PGP demonstration is to adjust current year expenditures for current year casemix. Therefore, a concurrent model is appropriate for the PGP demonstration.

The following is an example of how concurrent casemix adjustment works. Consider a hypothetical beneficiary with the following diagnoses:

Table 13-1

Hypothetical Example of Concurrent DCG-HCC Model Relative Risk Score for a Beneficiary

<u>Diagnostic Category</u>	<u>Label</u>	<u>Incremental Predicted Expenditures (\$1997)</u>
HCC 7	Metastatic Cancer	10,320
HCC 21	Protein-Calorie Malnutrition	10,038
HCC 31	Intestinal Obstruction/Perforation	7,337
HCC 74	Seizure Disorders and Convulsions	1,257
HCC 75	Coma	10,118
HCC 131	Renal Failure	8,865
Total Predicted Expenditures		47,935
Mean Expenditures, Medicare population		5,157
Relative Risk Score		9.30

SOURCE: Pope *et al.*, 2000, Table 7-2, Model 6.

Table 13-1 shows how an incremental predicted expenditure is associated with each diagnosis.³² The incremental predicted expenditures are summed to arrive at total predicted expenditures. Total predicted expenditures is then divided by mean expenditures in the 1997 model calibration sample to yield a relative risk score of 9.30. This risk score means that a beneficiary with these diagnoses is expected to incur Medicare expenditures 9.30 times more than the average Medicare beneficiary. Relative

³² The most expensive of related diagnoses are selected by the model. For example, if the beneficiary had another cancer diagnosis, that would receive a weight of zero because HCC 7 is at the top of the cancer diagnosis hierarchy.

risk scores greater than 1.00 indicate an expected costliness greater than the average, risk scores less than 1.00 an expected costliness less than the average.

The current version of the concurrent DCG-HCC model contains 89 diagnostic categories (Pope *et al.*, 2000, Table 7-2, Model 6). As the above example shows, these categories are combined additively to form a risk score. That is, an individual may be assigned to more than one diagnostic category, the weights of which are then combined into an overall risk score. The diagnostic categories are not mutually exclusive.

To apply risk adjustment to populations, the average of individual beneficiaries' risk adjustment scores are taken. For example, the average risk adjustment score for all beneficiaries assigned to a PGP might be 1.19. This indicates that, on average, PGP beneficiaries are expected to incur Medicare expenses 19% greater than average.

In the PGP demonstration, base year per capita expenditures for PGP beneficiaries and for comparison group beneficiaries will be adjusted for change in casemix from the base year to the performance year. This will be done by multiplying base year per capita expenditures by the ratio of average risk factors for the performance and base years. Consider the example in Table 13-2. We present hypothetical unadjusted and casemix-adjusted per capita expenditures and growth rates for a PGP and its comparison group. The comparison group consists of qualifying beneficiaries living in two market area counties A and B, weighted 80% and 20%, respectively.³³ Unadjusted, the growth rate in per capita expenditures is a high 6%. This exceeds the 2%

³³ The county weights are determined by the proportion of PGP beneficiaries residing in the two counties. See Recommendation 3 above for explanation and discussion. Typically, these weights will vary between the base and performance years. In this example, for simplicity, we have assumed that they are the same.

Table 13-2

Hypothetical Example of Casemix Adjustment of Expenditure Growth and Medicare Savings Calculation

	<u>Unadjusted Per Capita Expenditures</u>			<u>Risk Adjustment Factors¹</u>		<u>Casemix-adjusted</u>	
	<u>Base</u>	<u>Performance</u>	<u>Growth Rate</u>	<u>Base</u>	<u>Performance</u>	<u>Expenditures, Base²</u>	<u>Growth Rate³</u>
PGP beneficiaries	\$5,000	\$5,300	6.0%	1.00	1.10	\$5,500	-3.6%
Comparison group ⁴							
County A (weight = 0.8)	6,000	6,100	1.7%	1.18	1.19	6,051	0.8%
County B (weight = 0.2)	5,700	5,900	3.5%	0.89	0.87	5,572	5.9%
Total (weighted average)	5,940	6,060	2.0%	--	--	5,955	1.8%
<u>Medicare savings calculation:</u>			<u>Unadjusted</u>				<u>Casemix-Adjusted</u>
Target (PGP base expenditures*comparison growth rate)=			\$5,101				\$5,597
Actual PGP expenditures, performance year =			5,300				5,300
Per capita Medicare savings (target minus actual) =			-199				297
			Medicare savings is negative, no bonus is paid.				Medicare savings is positive, a bonus may be paid.

¹ Average expected costliness of health status casemix relative to Medicare national average of 1.00.

² Base year per capita expenditures adjusted for performance year casemix, i.e., multiplied by the ratio of performance year to base year casemix.

³ Percentage difference between performance year and adjusted base year per capita expenditures.

⁴ The comparison group consists of Medicare beneficiaries residing in two market area counties, A and B. The counties' per capita expenditures are weighted 80% and 20%, respectively, according to the shares of PGP beneficiaries residing in the two counties.

SOURCE: Health Economics Research, Inc.

Table 13-3

Simulated Bonus Payments With and Without Casemix Adjustment

Potential PGP Site	Growth Rate in State Per Capita Expenditures^{1,2}	Growth Rate in Site Per Capita Expenditures¹	Difference (RSPS)³	Bonus Payment to Site⁴	Medicare Program Savings⁵
I. Mean Bonus Payment Using 2% Threshold Without Case-mix Adjustment					
PGP A	6.1 %	14.9 %	-8.7 %	0.0 %	0.0 %
PGP B	6.1	15.1	-8.9	0.0	0.0
PGP C	6.5	4.9	1.6	0.0	1.6
PGP D	5.9	1.0	5.0	1.5	3.5
PGP E	8.3	3.7	4.6	1.4	3.2
PGP F	2.5	7.8	-5.3	0.0	0.0
PGP G	11.2	2.4	8.9	2.7	6.2
II. Mean Bonus Payment Using 2% Threshold With Case-mix Adjustment					
PGP A	3.9 %	8.9 %	-5.0 %	0.0 %	0.0 %
PGP B	3.9	6.9	-3.0	0.0	0.0
PGP C	1.9	1.5	0.4	0.0	0.4
PGP D	2.6	-3.5	6.1	1.8	4.3
PGP E	9.8	11.1	-1.3	0.0	0.0
PGP F	3.6	6.5	-2.9	0.0	0.0
PGP G	8.0	2.4	5.6	1.7	3.9

NOTES:

¹ Between 1993 and 1994.

² Each site's state proxies for its "market area."

³ Relative Savings Per Patient Seen (RSPS) is the difference in growth rates between the State and PGP Site.

⁴ Bonus payment to PGP Site (B) = RSPS × PCR × SR, where PCR is the Patient Capture Ratio

and SR is the sharing rate. In this analysis, we assumed that PCR = 0.4 and SR = .75.

⁵ Medicare Program savings = RSPS - B.

SOURCE: Adamache *et al.* (1998)

growth rate of the comparison group, so Medicare savings is negative and no bonus is paid. However, the PGP's predicted expenditures (casemix) grew 10% more expensive, so its casemix-adjusted expenditure growth is negative 3.6%. This is less than the comparison group's casemix-adjusted growth rate of 1.8%.³⁴ On a casemix-adjusted basis, per capita Medicare savings is \$297 or 5.6% of the PGP's performance year per capita expenditures. Subject to bonus thresholds, quality performance, sharing of savings with Medicare, and withholds, the PGP may be paid a performance bonus.

Health Economics Research, Inc. (Adamache *et al.*, 1998) simulated the impact of casemix adjustment on PGP demonstration bonus payments using actual historical data for a small sample of PGPs. The concurrent DCG-HCC model was used as the casemix adjuster. Table 13-3 is reproduced from Tables 4-1 and 4-2 in Adamache *et al.*, 1998. HER found that casemix adjustment had a significant impact on expenditure growth rates. It tended to lower measured growth rates of both simulated sites and of associated market areas, and to lessen the variability in site relative to market area growth rates. The range across sites in difference between state (a proxy for market area) and site growth rates fell from -8.9% to 8.9% without casemix adjustment to -5.0% to 6.1% with casemix adjustment (compare the "difference" column in panels I and II of Table 13-3). One simulated site, PGP E, would have been paid a bonus without casemix adjustment, but not with casemix adjustment. Casemix adjustment reduced one site's expenditure growth rate from 15.1% to 6.9% (PGP B), and another's (PGP A) from 14.9% to 8.9%. It

³⁴ Each comparison group county's (A's and B's) base year expenditures are multiplied by its ratio of performance to base year risk factors. Then a weighted average of these adjusted base year expenditures is taken and compared to the weighted average performance year expenditures to determine the casemix-adjusted growth rate of 1.8%.

increased a third's (PGP E) from 3.7% to 11.1%. We conclude that casemix adjustment is desirable to reduce fluctuations in expenditure growth rates related to casemix changes. We also conclude that casemix adjustment may substantially impact bonus pool calculations for the PGP demonstration.

Although the basic concurrent DCG-HCC risk adjustment model has already been developed, some extensions/refinements will be necessary to apply it to this demonstration. CMS reserves the right to make changes to the casemix adjustment model for the PGP demonstration. We anticipate that the following refinements will be needed and made. First, the case mix model will need to be reestimated including beneficiaries eligible for Medicare because of End Stage Renal Disease (ESRD). They were not included in M+C models because ESRD beneficiaries are not currently eligible to enroll in M+C plans. But ESRD beneficiaries will be involved in the PGP demonstration. Second, the model will need to be reestimated for users of the specified evaluation and management services used to assign beneficiaries to PGPs (see Recommendation 2). The model currently is estimated on all Medicare beneficiaries, both users and non-users of evaluation and management services. Third, the model will need to be reestimated top-coding expenditures at \$100,000 to be consistent with the outlier policy for the PGP demonstration. Fourth, the model will need to be reestimated including major organ transplant procedures. Fifth, the diagnostic categories included in the model will need to be reviewed to eliminate those that may be very sensitive to diagnostic coding changes or variations. Sixth, the model will need to be recalibrated on more recent 1999/2000 data. Seventh, given these changes, some adjustments may be

necessary to ensure that the model is calibrated correctly for different Medicare subgroups.

Customizing the concurrent DCG-HCC risk adjuster for this demonstration represents a significant task for the demonstration technical support contractor, HER. But HER is also a CMS contractor for developing Medicare risk adjustment models, and thus is very familiar with estimating these models. We expect that the model will only need to be customized/updated/reestimated once for this demonstration, using Medicare datafiles developed in HER's risk adjustment work for CMS. This can be done during the base year of the demonstration using the most recent risk adjustment datafiles then available. By completing this work during the base year, scheduling conflicts with the technical support contractor's work in computing bonuses, etc. in demonstration performance years can be avoided.

Recommendation 14: Adjusting Performance Targets for Effects of Medicare Payment Policy

- Medicare payment policy changes will be reflected in the expenditures of both the participating PGP and its comparison group. Hence, Medicare savings and PGP bonuses under the demonstration will not be sensitive to changes in Medicare payment policies. No special adjustments for changes in payment policy are recommended.

Background: A participant's performance target for a performance year during the demonstration is derived by multiplying the base amount – a site's adjusted base year per capita expenditures – by the applicable expected growth rate. The expected growth rate for a participant is the ratio of the comparison population's per capita expenditures in the performance year to the base year. The comparison population's per capita

expenditures in the performance year will reflect changes in Medicare payment policies. The expected growth rate, thus, implicitly adjusts the participant's base amount for all changes in Medicare payment rates and policies.

Discussion: Major and minor changes in payment policies are a constant feature of the Medicare program. Major changes are usually the result of Congressional action while minor changes are instituted to refine policies. In addition, payment rates for inpatient services change annually as a consequence to updates in the standardized amounts, DRG weights, and the area wage index for PPS hospital inpatient discharges and similar updates for physician payments, outpatient payments, and so forth.

All providers covered by specific policies are affected when the policy changes or its payment rates are updated. Hence, we expect that policy changes will have similar effects on both the PGP participants and their comparison populations. The *differential* effects of payment policy will probably be small. Most markets that have eligible PGPs contain a wide spectrum of providers and have multiple providers of specific services. With this range of providers in the comparison group, any differential effects arising from payment policy per se will probably be small. While the impact of policy changes on individual providers might vary, subgroups of providers will experience similar average effects. Even if the PGP itself provides a very concentrated or unusual set of services, bonus calculations under the demonstration are based on total per capita expenditures for assigned beneficiaries, not just those provided by the participating PGP. Thus, for any differential effect to arise, assigned beneficiaries would have to receive a very different set of services than comparison beneficiaries on average, which would be highly unusual for two large groups (15,000 or more) of beneficiaries in the same market

area.

Attempting to adjust PGP demonstration calculations for annual changes in Medicare payment policy is administratively infeasible. Fortunately, the adjustment is made implicitly in the design of the demonstration. This occurs through the use of a comparison group that is affected by the same payment policy changes as the participating organization.

Part V

Quality Targets and Bonuses

V. Quality Targets and Bonuses

Providing direct financial incentives for quality improvement is one of the innovative features of the PGP demonstration. As a result, however, it is expected to be complex and potentially sensitive. Nonetheless, direct quality incentives for providers have been implemented by a range of business health care purchasing coalitions and health plans in recent years, so there are a number of precedents (Fraser *et al.*, 1999; Integrated Healthcare Association, 2002; IOM, 2001a). Moreover, the IOM (2001a) recently recommended that direct incentives for quality improvement be applied more widely by both public and private purchasers, and specifically recommended that CMS develop pilot tests and evaluations of this approach and others aimed at quality improvement. The PGP demonstration thus provides a fortuitous opportunity to explore and test the types of incentive-based quality improvement methods highlighted as priorities by the IOM.

To provide for greater acceptability by PGPs, we have focused on quality measures that have been validated in previous studies or are currently in use, are relatively non-controversial, and avoid placing excessive data collection or administrative burdens on participating PGPs. We have included several measures used by NCQA in its HEDIS program, as well as others applied federal quality initiatives, including Medicare's Health Care Quality Improvement Program (HCQIP) and the Agency for Healthcare Research and Quality's (AHRQ) Healthcare Cost and Utilization Project (HCUP) Quality Indicators. We also emphasize the use of quality measures

which can be expected to lead to lower costs for the Medicare program as quality is improved.

For Part V we have developed recommendations regarding six separate quality-related issues. As a result, our discussion of procedures for developing quality-related targets and incentives is presented in six sections, Recommendations 15-20.

Recommendation 15: General Approach for Quality Indicators

- HER recommends the demonstration include 8 process and outcome quality indicators.
- CMS will work with participating PGPs to identify the final set of indicators.
- CMS will retain the right to make the final selection of indicators.

We believe a relatively small number of indicators (less than 10) will have several advantages. First, it reduces the complexity and cost of collecting data and calculating performance comparisons and incentive payments. Second, a smaller number of indicators can be focused on high volume clinical conditions which affect a large percentage of the Medicare population, or on conditions known to be associated with frequent or significant quality problems. Third, focusing on higher volume conditions will provide larger sample sizes to improve the statistical reliability of comparisons between quality targets and actual PGP performance.

The number of indicators could be increased in the future if this new PGP reimbursement system is implemented more widely by Medicare after the demonstration is completed. However, HER recommends that the short-run goal for the demonstration should be to focus on “pilot testing” a smaller set of indicators.

The main disadvantage of a smaller number of indicators is the limited scope of diseases covered. In theory, PGPs could respond by focusing their quality improvement efforts on the conditions being measured, while neglecting other conditions not captured by the quality indicator data. However, similar indicators can be combined into an index or aggregate indicator, so that one indicator covers multiple diseases. That approach is applied in two of the specific indicators recommended below.

A potential advantage of using a larger number of indicators would be improved comprehensiveness. However, additional indicators would need to be identified that are broadly considered acceptable, valid, and reliable. Moreover, the complexity of a system with many indicators would be more difficult to implement and could dilute the incentive impact of quality improvement efforts targeted at specific conditions.

Recommendation 16: Selecting Specific Process and Outcome Quality Indicators

- The major focus of the demonstration will be on measuring process indicators of quality. They are the indicators most easily measured and most relevant to the medical care operations of PGPs.
- Selected outcome indicators will also be included.

Quality of care can be measured and analyzed using several different concepts. The basic choices are: (i) structure (or inputs); (ii) process; and (iii) outcomes (including patient satisfaction). The BIPA 2000 legislation mandates a focus on process and outcome indicators for the PGP demonstration.

Process indicators are more commonly used in health care quality improvement programs since they are more directly under the control of medical providers. Moreover, process indicators generally provide for annual or more frequent data collection, which

will facilitate the annual monitoring efforts required for the PGP demonstration. Outcome indicators are sometimes used in quality improvement efforts, but they are often less practical to measure and analyze. Nonetheless, outcomes are what beneficiaries care about in the end, so we recommend including selected outcome indicators, focusing on those where the measurement issues are less of a concern.

We anticipate the specific process and outcome indicators used in the demonstration will be selected from among the following indicators recommended by HER:

Recommended Process Indicators

- Eye examination every 2 years for beneficiaries with diabetes.
- Hemoglobin A1c test every year for beneficiaries with diabetes.
- Lipid profile test every 2 years for beneficiaries with diabetes.
- Mammogram every 2 years for female beneficiaries between the ages of 52-69.
- Both chest radiograph and electrocardiogram ≤ 3 months after initial diagnosis of congestive heart failure (CHF).
- Left ventricular ejection fraction (LVEF) test conducted during the current year for beneficiaries hospitalized with a principal diagnosis of CHF during the current year.
- One physician visit every 6 months for beneficiaries with any one of the following 4 chronic disease diagnoses: chronic stable angina; chronic obstructive pulmonary disease (COPD); CHF; or diabetes.
- Annual influenza vaccinations for all beneficiaries age 65 or older.
- Pneumonia vaccination status for all beneficiaries age 65 or older (ever had a pneumonia vaccination).

Recommended Outcome Indicators

- Number of hospital admissions per year for three high volume ambulatory care sensitive conditions (ACSCs) per 1000 Medicare

beneficiaries. The conditions include asthma/COPD, CHF, and bacterial pneumonia.

- Patient satisfaction maintained above a lower bound level (specific survey instrument and target level to be determined).

We considered four factors in selecting specific quality indicators to recommend for the PGP demonstration. First, focusing on indicators already in widespread use in national quality improvement efforts, such as the Health Plan Employer Data & Information Set (HEDIS) and Medicare's Health Care Quality Improvement Program (HCQIP). The Medicare HCQIP provided the indicators for biennial eye exams, annual hemoglobin A1c tests, and biennial lipid profile tests for diabetics (CMS, 2001a; Jencks *et al.*, 2000). HEDIS provided the indicators for biennial mammograms, influenza vaccinations, and pneumonia vaccinations (NCQA, 2001). The ACSC measure was developed under a CMS contract for application to the Medicare population (McCall *et al.*, 2001). Similar measures have been developed for AHRQ's HCUP quality indicators (Davies *et al.*, 2001). Both types of patient satisfaction surveys considered for use in this demonstration (CAHPS[®] and AMGA, discussed below) are currently in use in national quality improvement efforts (Goldstein *et al.*, 2001; Sanderson-Austin & Wetzler, 2001). The LVEF indicator is being used by CMS for incentive payments to Medicare+Choice plans (CMS, 2001). The other recommended indicators were developed in a study which focused specifically on developing claims-based quality indicators for the Medicare program (Asch *et al.*, 2000).

Second, including indicators that have been developed for – or applied to – the Medicare population, and published in peer-reviewed literature. Each recommended indicator has been calculated for Medicare study populations using Medicare claims data

or patient surveys and published in peer-reviewed literature (Asch *et al.*, 2000; Goldstein *et al.*, 2001; Jencks *et al.*, 2000; McCall *et al.*, 2001; Wu and Pope, 2002). Thus the recommended indicators should have credibility with participating PGPs.

Third, including indicators that represent a range of different types of medical care interventions. Our recommended indicators provide a relatively broad perspective on a PGP's quality of care, especially for a list limited to about 10 indicators overall.

Fourth, focusing on indicators that apply to high volume diseases or clinical conditions to ensure sufficient sample sizes for reliably evaluating the performance of PGPs and face validity for the indicator's salience to participating PGPs. Each recommended indicator applies to a significant number of Medicare beneficiaries.

The importance of eye examinations and hemoglobin A1c tests are well established for diabetics and have been validated by randomized controlled trials (Asch *et al.*, 2000; Jencks *et al.*, 2000; NCQA, 2002). Eye exams enable early detection of retinal damage and may reduce long-run costs by preventing or delaying this frequent complication of diabetes. Eye exams for diabetics is a quality indicator also used in an earlier Health Economics Research (HER) project evaluating potential methods for monitoring quality of care for fee-for-service Medicare beneficiaries (McCall *et al.*, 1998). Jencks *et al.* (2000) found that only 69 percent of diabetic FFS Medicare beneficiaries were receiving care which met the biennial eye exam guideline in the median state, in their study which included all 50 states and the District of Columbia and Puerto Rico.

Hemoglobin tests provide for better control of diabetes and may help prevent or delay a range of complications. Jencks *et al.* (2000) found that only 71 percent of diabetic FFS Medicare beneficiaries were receiving care which met the annual test guideline in the median state.

Lipid profile tests are important for monitoring the risk of cardiovascular disease, which afflicts many diabetics (CMS, 2001a). High lipid levels are a modifiable risk factor and thus careful monitoring can lead to effective interventions. Jencks *et al.*, (2000) found that only 57 percent of diabetic FFS Medicare beneficiaries were receiving care which met the biennial test guideline in the median state.

Biennial mammograms for women aged 52-69 have been identified as a useful quality indicator based on randomized controlled trials (Asch *et al.*, 2000). Some debate does continue on the appropriate ages to include for this indicator, however. Asch *et al.*, (2000) included women up to age 75 (their sample was restricted to beneficiaries age 65 or more). However, both HEDIS and Medicare's HCQIP uses the age range of 52-69, so we recommend that age range for this demonstration (HEDIS, 1997; CMS, 2001b). Jencks *et al.*, (2000) studied this indicator in their recent national analysis of claims data for Medicare FFS beneficiaries. They found that only 56 percent of women of those ages received mammograms in the median state during the two-year period covering 1997 and 1998.

The quality indicator for chest radiographs and electrocardiograms in the first three months after an initial diagnosis of CHF was tested by Asch *et al.*, (2000) in their recent study of claims-based indicators for Medicare. They found that only 64 percent

and 68 percent, respectively, of eligible beneficiaries with CHF received those tests during the indicated time period.

The LVEF test indicator for beneficiaries hospitalized for CHF is currently being used by CMS to provide incentive payments to Medicare+Choice plans (CMS, 2001). It is based on recommendations by the American College of Cardiology, the American Heart Association, and a CMS clinical panel which studied quality indicators for CHF (Wu and Pope, 2002). Jencks *et al.* (2000) found that only 65 percent of FFS Medicare beneficiaries were receiving care which met this guideline in the median state. However, their study was limited to inpatient records sampled over a six-month period. Wu and Pope (2002), found a higher rate of compliance with this guideline, 79 percent, through analysis of a two-year sample of inpatient claims. The annual indicator recommended here is adapted from the approach taken by CMS, in order to focus on a single year measurement period.

The next quality indicator measures the frequency of physician visits for chronically ill Medicare beneficiaries. It indicates that frequent visits (at least one visit every 6 months) are appropriate for beneficiaries with any one of four common chronic diseases (chronic stable angina, COPD, CHF, and diabetes). Other diseases could presumably be added to this list, but this indicator focuses on diseases studied in the peer-reviewed literature and those with high incidence in the Medicare population, to provide for larger sample sizes. Asch *et al.*, (2000) studied this indicator for these diseases and found it met for each disease for more than 90 percent of the Medicare FFS beneficiaries in their sample. Of 20,994 patients they studied with chronic stable angina, they found

that 96 percent had one physician visit at least every six months. For 38,947 patients with chronic obstructive pulmonary disease, they found that 93 percent had at least one physician visit every six months. For 35,858 patients with congestive heart failure, they found that 96 percent had at least one physician visit every six months. For 47,841 patients with diabetes, they found that 93 percent had at least one physician visit every six months.

This indicator is designed to provide a target which will be relatively easy for PGPs to meet if undertreatment is not a strategy used to reap increased bonus payments through the cost efficiency incentives available in this demonstration. However, it is intended to detect undertreatment if it exists. Incentives for undertreatment are inherent to some extent in all reimbursement systems which promote lower utilization, whether in this demonstration or, more strongly, under capitated reimbursement.

Influenza and pneumonia vaccination rates are preventive care process indicators, which are included in both HEDIS and the CDC's Behavioral Risk Factor Surveillance System (BRFSS). Jencks *et al.*, (2000) analyzed these indicators using BRFSS data in their recent national study on quality of care for Medicare FFS beneficiaries. They found that only 66 percent of beneficiaries age 65 or over had received an annual flu shot in the median state. They also found that only 46 percent of beneficiaries age 65 or over had ever received a pneumonia vaccination in the median state. For the PGP demonstration, these indicators would be measured using patient surveys. That is the approach used by HEDIS, which conducts surveys to measure these indicators using the CAHPS[®]. The

CAHPS[®] survey is described in more detail below in the discussion of the patient satisfaction indicator.

The first recommended outcome indicator is reducing hospital admissions for ambulatory care sensitive conditions (ACSCs). It has the potential to reduce costs at the same time as it focuses attention on the quality of primary care for chronic diseases. HER recently conducted a study which involved developing and testing a set of 15 ACSCs specifically for Medicare beneficiaries (McCall *et al.*, 2001). Those ACSCs were all defined using the primary diagnosis code for inpatient admissions. A list of the 15 conditions is included below in Appendix D.

HER recommends that the ACSC admissions measure used for this demonstration include analysis of data for the 3 highest-volume conditions from this group of 15: asthma/COPD, CHF, and bacterial pneumonia. They represent about 70 percent of all ACSC admissions for the overall list of 15 conditions (McCall *et al.*, 2001). Each of these three high volume conditions was also included in the list of ACSCs developed for AHRQ's recent update and refinement of its HCUP quality indicators (Davies *et al.*, 2001).

ACSCs are considered an “outcome” indicator since avoiding unnecessary admissions means reduced morbidity for the patient, and probably reduced mortality in some cases as well. Quality of life is presumably also enhanced when unnecessary hospital admissions are avoided.

ACSCs are medical conditions for which physicians broadly concur that a substantial proportion of cases should not advance to the point where hospitalization is

needed if they are treated in a timely fashion with adequate primary care and managed properly on an outpatient basis (McCall *et al.*, 2001). Because lack of adequate primary and outpatient care does, in fact, often result in preventable admission for these patients, the annual rate of admissions per 1000 beneficiaries provides a practical way of identifying priority areas for improving quality and better coordinating care.

Reducing ACSC admissions also has an additional benefit for PGPs under the demonstration. By directly reducing utilization and cost, they will also aid PGPs in meeting their cost targets. Most of the other recommended quality indicators will probably also reduce costs at some point, for example by reducing complications and comorbidities for beneficiaries with chronic illnesses, but those effects may not be experienced in the same year in which the additional expenses are incurred for meeting the quality guideline.

We recommend creating an index of ACSC admissions for these three conditions by simply adding admissions for these conditions together, and then calculating a rate of ACSC admissions per 1,000 Medicare beneficiaries for each participating PGP. Adding the admissions for each condition means that the index will be “self-weighting,” that is, the more common conditions among these three will have larger weights in the index.

The HER study on ACSCs analyzed hospital admission data for over 4 million Medicare+Choice (M+C) enrollees, who experienced 191,323 hospitalizations between July 1997 and June 1998 (McCall *et al.*, 2001). They experienced an overall rate of 47.2 ACSC admissions per 1,000 Medicare beneficiaries. This rate included 14.2 ACSC admissions for congestive heart failure, 10.7 admissions for pneumonia, and 8.4

admissions for asthma/COPD. Thus these three high volume conditions represented a total of 33.3 ACSC admissions per 1,000 Medicare beneficiaries, or 70.1 percent of the overall rate.

The sum of these three high volume conditions represents an ACSC index which has been recommended by HER to CMS in a more recent ACSC study for monitoring quality of care for M+C plans (Pope *et al.*, 2002). HER studied 14 different possible ACSC indexes constructed from the overall list of 15 ACSC conditions. The index constructed using the three highest-volume conditions was found in that study to have more favorable characteristics than the other indexes across several criteria, including reliability, parsimony, and interpretability, and similar characteristics for validity and administrative feasibility.

The final indicator HER recommends is a second outcome indicator. It will measure patient satisfaction using a survey-based approach. At present, selection of a recommended survey instrument is under consideration. We are currently reviewing two main options. The first option is the new physician group version of the Consumer Assessment of Health Plans (CAHPS[®]) survey, called the Provider Group CAHPS[®] (AHRQ, 2001). Versions of the CAHPS[®] survey have been developed for Medicare Managed Care and fee-for-service programs (Goldstein, 2001). It has already been applied extensively for studying consumer satisfaction in M+C plans (AHRQ, 1999; Goldstein *et al.*, 2001). The new Provider Group CAHPS[®] is similar to the health plan version, and it has been extensively field tested, but it is not slated to be finalized and released until later this year (AHRQ, 2001). The CAHPS[®] approach also has the benefit

of including survey measures for our recommended influenza and pneumonia vaccination indicators in its applications for HEDIS.

The second option for measuring patient satisfaction is an existing survey which is currently being conducted by a number of mainly larger PGPs, the AMGA Patient Satisfaction Benchmarking Program (AMGA, 2002; Sanderson-Austin & Wetzler, 2001). Its survey instrument is shorter than the Provider Group CAHPS[®] (13 items versus 43 items). However, the AMGA survey includes all patients seen by a PGP, and hence has not been targeted to Medicare patients. An available option for adding several additional questions could enable this survey to identify Medicare beneficiaries and to capture data on influenza and pneumonia immunization rates.

A further difference is that the AMGA survey is designed to be completed on site, at the PGP, at the conclusion of a physician visit (AMGA, 2002). The patient focuses on evaluating the characteristics of the particular visit just concluded, and drops off the completed questionnaire prior to leaving the PGP. The goal is to conduct at least 30 surveys for each provider in a PGP, and then submit all of the surveys for analysis and scoring. Analysis can focus on the PGP overall, on the different specialty groups within the PGP, or on individual providers. In contrast, the Provider Group CAHPS[®] is designed to be completed by mail or telephone (or both) when the patient is at home. Only a random sample of patients are selected. (For Medicare Managed Care plans up to 600 beneficiaries are surveyed for each plan.) Patients are asked to recall their experiences with the PGP over the past 12 months for most questions. Analysis focuses on the PGP overall.

Further consideration is needed to select the best approach for measuring and monitoring patient satisfaction for the PGP demonstration. However, the Provider Groups CAHPS[®] and AMGA Patient Satisfaction Benchmarking Program indicate that valid and reliable survey instruments are available for this purpose. It will be feasible to measure patient satisfaction; the main questions will be to select the best approach for the purposes of this demonstration and to identify the additional budget requirement needed for implementing the survey.

Either alternative for measuring patient satisfaction (and the survey-based measures for influenza and pneumonia vaccination rates) will require additional budget resources for the PGP demonstration for data collection and analysis. The other quality indicators recommended above are measured using claims data, which will already be collected for analyzing the cost performance of participating PGPs. The patient satisfaction and vaccination measures will require a separate survey data collection process. Estimating the budget requirement for this process will require selecting a preferred option for the survey, identifying the number of beneficiaries to be surveyed for each PGP, and identifying a survey research vendor to administer the data collection effort.

Developing specific targets for the patient satisfaction indicator will require additional analysis after a preferred survey instrument has been selected. However, standard summary measures have been developed for both the Provider Group CAHPS[®] and the AMGA Patient Satisfaction Survey, so developing specific targets should be feasible.

For example, if the Provider Group CAHPS[®] is selected, then the targets could be based on the four global ratings of patient satisfaction that are included in the survey (AHRQ, 2001; Goldstein *et al.*, 2001). Each of these provides a rating of one aspect of patient satisfaction on a 0-10 scale. Data for these ratings for PGPs participating in the demonstration could be compared with external standards, or averaged to create a single measure and then compared with an external standard. CAHPS[®] also provides composite satisfaction measures on getting needed care, getting care quickly, doctor communication, office staff, and paperwork and customer service (AHRQ, 2001; Goldstein *et al.*, 2001). The AMGA Patient Satisfaction Survey focuses on individual items covering waiting time, convenience, telephone accessibility, technical and interpersonal skills of the provider, time spent with the provider, and an overall assessment of the visit (AMGA, 2002). AMGA maintains a database of survey responses for benchmarking; over one million surveys were included in the database as of July, 2000 (Sanderson-Austin & Wetzler, 2001). Both the CAHPS[®] and AMGA surveys provide data on the respondent's age, gender, and health status, which can be used to risk-adjust patient satisfaction ratings (AHRQ, 1999; Zaslavsky *et al.*, 2001).

The primary goal for the patient satisfaction measure is to guard against the risk of undertreatment in response to the demonstration incentives, so the target could be set at the average level for physician groups on these measures. Or possibly even slightly below the average if the goal is to identify those PGPs falling to very low levels, and to provide for incentive payments for all of those PGPs doing better than a “lower bound”.

An additional issue for measuring patient satisfaction and the other survey-based measures for the PGP demonstration will be identifying the sampling frame for survey data collection. Since the demonstration focuses on assigned beneficiaries, defined by claims data for E&M services provided to beneficiaries by PGPs (Recommendation 2), the assignment process cannot be conducted until claims data are available, after at least a 6-8 month lag time. As a result, the patient satisfaction surveys will need to be administered to patients assigned to a PGP in a prior year.

For example, in December of the demonstration's first performance year, only claims data for the base (prior) year will be available. We can identify patients assigned to a PGP in the base year at that point. The patient satisfaction and vaccination rate survey for the first performance year could then be conducted in January and February of the second performance year, if the survey is of the type which asks the patient to recall their health care experience over the previous 12 months. Using base year assigned beneficiaries will have the disadvantage of missing new patients receiving treatment from the PGP for the first time in the performance year. However, it will have the advantage of including patients who have chosen to discontinue receiving treatment from the PGP, which presumably will include a number who have discontinued treatment due to low levels of patient satisfaction.

Other claims-based process and outcome indicators are possible, as shown in Appendix C. They could also be considered for use in the demonstration. The eight claims-based indicators recommended here (all of the recommended ones except the patient satisfaction and influenza and pneumonia vaccination indicators) are intended to

limit the overall number of indicators measured. At the same time, they include two aggregate indicators, covering multiple diseases (admissions for three types of ACSCs and frequency of physician visits for four chronic conditions), which provide for coverage of a fairly broad range of medical conditions.

Each of the eight claims-based quality measures will be calculated using specifications provided by HCQIP, HEDIS, or other published sources. Those sources provide specifications for the numerators and denominators for each indicator. They also provide lists of the specific diagnosis and procedure codes used to calculate each indicator.

The statistical properties of the recommended claims-based quality indicators will be further studied during a simulation of the PGP demonstration which will be conducted over the next several months using Medicare claims data from 1997 through 2000 for a sample of 9 PGPs. This analysis will examine variations in the quality indicators across PGPs and over time. This analysis will be used to further refine the methods and targets applied for the quality indicators, as appropriate. It will also be used to evaluate the extent to which risk adjustments should be applied for these quality indicators for purposes of this demonstration.

Finally, in order to provide for innovation in quality indicators, we recommend that PGPs be eligible to propose up to two alternate quality measures on their own initiative. They could be substituted for two of the recommended quality measures specified in advance by CMS. Specification of the indicators eligible for substitution is

intended to ensure that all participating PGPs will be required to meet a core set of at least six indicators.

For this substitution policy, we also recommend that alternate quality indicators proposed by PGPs be reviewed and approved by CMS prior to adoption for the PGP demonstration. In general, measures proposed by PGPs should be considered if they are measurable using claims data, they apply to diseases affecting a significant proportion of the Medicare beneficiary population, and they include targets or benchmarks that can be externally verified. CMS should reserve the right to accept or reject proposed alternate indicators. In the absence of alternate indicators that are proposed by PGPs and accepted by CMS, all participating PGPs should be required to meet quality targets for the final set of eight indicators identified for the demonstration.

While allowing PGPs to propose up to two alternate indicators provides flexibility and promotes innovation, it also has two potential disadvantages. First, participating PGPs may propose indicators on which they expect to do well. This could result in biased measurement of their quality of care. Hence CMS should carefully scrutinize any alternate indicators proposed by PGPs. Second, allowing alternate indicators will result in some inconsistencies in measurement of quality across the PGP sites participating in the demonstration. This may be perceived as inequitable and may limit, to some extent, the validity of cross-site comparisons. However, by maintaining a common set of at least six indicators measured for all PGPs, this concern should be mitigated.

Recommendation 17: Data Sources for Measuring Quality Indicators

- Administrative billing ("claims") data will be the primary data source for measuring quality indicators.
- Patient survey data will also be collected to measure patient satisfaction and influenza and pneumonia vaccination rates.

Claims are low cost and take advantage of the central role of claims data in the PGP demonstration for calculating cost targets, cost performance comparisons, and Medicare Savings. As noted, 8 of the 11 recommended indicators can be measured and monitored using claims data. We also recommend patient surveys since they enable collection of data not available in claims, including patient satisfaction and influenza and pneumonia vaccination rates.

Medicare claims have the advantage of including reasonably complete data. Since claims are primarily used for billing, providers have a direct financial incentive to ensure all possible bills are submitted for reimbursement. Moreover, the necessary Medicare claims files will already be collected for the PGP demonstration for calculating the cost-related targets and incentives. Thus using claims data for measuring quality means there will be no added data collection or administrative burdens for the participating PGPs or for the demonstration's administrators.

The main disadvantages of Medicare claims data include a lack of pharmaceutical data, limited clinical information, and a lack of patient self-reports on satisfaction with care and quality of life. However, several recent studies have identified, tested, and evaluated a wide range of quality indicators which can be measured and tracked using Medicare claims data, which means there are a number of validated, claims-based

indicators already developed and available for use in the PGP demonstration (Asch *et al.*, 2000; CMS, 2001; Garnick *et al.*, 1994; Jencks *et al.*, 2000; Johantgen *et al.*, 1998; McCall *et al.*, 1998; McCall *et al.*, 2001; NCQA, 2002; Weiner *et al.*, 1995; Wu and Pope, 2002). As noted, a number of these indicators are summarized in Appendix C.

Alternate data sources may offer different types of quality-related information, but are generally more costly or are inconsistent across PGPs. For example, internal PGP information systems may have quality-oriented modules, especially if PGPs have ongoing internal quality monitoring efforts. They may have been developed in response to JCAHO or HEDIS initiatives, for example. They might include a range of useful clinical information, but are probably not standardized across PGPs, and thus could not be applied uniformly across the PGPs participating in the demonstration.

Medical record abstracts are often used for quality indicators where pharmaceutical data or detailed clinical information are needed. For example, a Medicare HCQIP indicator for acute myocardial infarction (AMI) patients measures whether or not patients received a prescription for a beta-blocker at discharge for those hospitalized for AMI (CMS, 2000). That indicator is measured using medical records data, since claims data do not record that type of information. However, medical record abstracts are costly. Given the limited number of indicators recommended for the PGP demonstration, we believe they are unnecessary, since there are a sufficient number of well-established claims-based indicators available. Claims data provide some measures of clinical issues, focusing on diagnoses and utilization of particular tests or services.

Patient surveys are also costly. As noted, implementing them will require additional funding for the PGP demonstration. However, they can provide information on self-reported patient satisfaction and health-related quality of life (HRQOL), and on influenza and pneumonia vaccination rates, which cannot be collected otherwise. As a result, we have recommended collection of patient survey data for this demonstration, to supplement the primary focus on claims-based quality indicators.

By recommending surveys on patient satisfaction we include some self-reported data for measuring the patient's own experience of health care. Collecting additional survey data on HRQOL would provide another type of self-reported information, but would add significantly to the length of the survey and add to the expense and complexity of the demonstration. With the limited number of indicators to be measured for this demonstration we believe the patient satisfaction data will be sufficient. We view patient satisfaction data as especially important for the PGP demonstration given the incentives for undertreatment that are inherent in its cost efficiency incentives.

Recommendation 18: Quality Targets for PGPs

- The demonstration will include two types of targets a PGP can meet in order to earn the quality portion of the PGP Bonus Pool:
 1. Achieving a pre-determined threshold level for a quality indicator.
 2. Demonstrating substantial quality improvement over time.
- Only a threshold target is recommended for patient satisfaction, since that indicator is primarily intended to provide a lower bound for detecting undertreatment.
- *Either* threshold or quality improvement targets could be met to earn a quality bonus for all of the other indicators.

An advantage of pre-determined, fixed thresholds is that they reward groups for quality achievements involving evidence-based goals. A disadvantage of the threshold approach is that it may discourage some PGPs from applying for the demonstration if the thresholds are set at very high levels (e.g., requiring 95 percent of diabetics to receive hemoglobin A1c tests each year), which might be hard for any PGP to achieve.

An advantage of setting targets based on substantial quality improvement over time is that they are consistent with the goals of continuous quality improvement (CQI) or total quality management (TQM) programs, which underlie many of the recent efforts toward quality improvement in the health sector (IOM, 2001a). This perspective recognizes that achieving “ideal” thresholds is often very difficult in actual medical care practice settings, given limits on overall resources due to reimbursement limitations, increasing limits on physician time, the large and rapidly growing number of quality guidelines being published and debated, limitations of available allied health staff and support staff, limitations of clinical and information systems, the growing independence of patient decision-making, and other factors. Instead of penalizing physicians or PGPs for failing to reach ambitious thresholds that may be somewhat beyond their control (at least in the short run), this approach rewards the realistic quality improvements that are possible in the short run (Chassin, 1996).

Quality assurance systems that focus on penalizing outliers (usually defined using thresholds) have often produce large amounts of wasteful, defensive behavior by physicians and other staff. The CQI or TQM approach focuses on “moving the mean” or improving average quality performance, rather than identifying and punishing outliers.

However, a disadvantage of quality targets focused on demonstrating improvement over time is that they may penalize groups that have already achieved high levels of quality. Those PGPs may find it much harder to show improvement than other groups starting from a base of much lower quality.

As a result, our recommended approach allows PGPs to earn quality bonuses in either of the two ways, through achieving thresholds or demonstrating improvement over time. This avoids the problems involved in using either method exclusively.

We recommend the following protocols be used to determine if a PGP has met either the pre-determined, fixed threshold or substantial improvement over time criterion for each indicator:

Pre-determined Threshold Targets

- The fixed threshold level of quality will be set at 75 percent or more of a PGP's eligible beneficiaries receiving the care specified by the indicator for six of the eight claims-based indicators (all except the indicators for ACSCs and for visits every 6 months for beneficiaries with the four specified chronic conditions), and for the two survey-based vaccination measures.

For example, if 75 percent of a PGP's diabetics had hemoglobin A1c tests in a given year, it would have met the quality target. Further improvement would not be necessary to continue to earn the quality bonus for that indicator. The threshold is set somewhat below 100 percent in recognition that perfect performance is usually not achievable, due to limited budgetary and other resources, imperfect internal PGP clinical and management systems, some degree of patient non-compliance with physicians' recommendations, and other factors. The 75 percent threshold is, however, significantly

above the national average rates for each of these indicators. As noted, studies by Asch *et al.* (2000) and Jencks *et al.* (2000) found rates ranging from 43 percent to 69 percent for these indicators in national studies using Medicare claims data. The target rate for the LVEF indicator was recently set at 75 percent by CMS for its incentives applied to Medicare+Choice plans (CMS, 2001).

- The threshold level will be set at 90 percent for physician visits every six months for beneficiaries with chronic stable angina, COPD, CHF, or diabetes.

This threshold is higher since the study by Asch *et al.* (2000) found that six-month visit rates were already above 90 percent for each of these chronic diseases. The threshold is not set even higher since this indicator is primarily intended to detect undertreatment if it occurs.

- The threshold level for meeting the quality target for the ACSC indicator will be set at the FFS national average rate per 1,000 Medicare beneficiaries.

For example, if the national average rate for a performance year is found to be 33.3 ACSC admissions per 1,000 Medicare beneficiaries, as was found in an HER study for the Medicare+Choice population (McCall *et al.*, 2001), then any rate below that figure would mean the PGP had met this target. (In contrast to the other quality indicators, where higher results are better, for ACSCs lower rates represent better performance.)

There are two reasons for using the national average figure as the threshold for the ACSC quality indicator. First, excessively low rates of ACSCs may indicate access problems and not quality improvement. Thus while a 100 percent performance target

may be desirable for other quality indicators, such as eye exams for diabetics, a less extreme approach is warranted for the ACSC indicator. Second, even this modest, national average threshold will save money for Medicare while at the same time encouraging improvements in quality. Reductions in ACSC admissions will directly and immediately benefit CMS due to the relatively high costs associated with hospital admissions. Improvements in other quality indicators may reduce costs for CMS in the medium term or long term, such as by reducing rates of complications suffered by diabetics, but they will not provide the same type of immediate cost savings as improvements in the ACSC indicator. At the same time, those forgone costs will also aid PGPs in meeting their cost targets under this demonstration.

- A lower-bound patient satisfaction threshold will be determined which will serve to detect undertreatment if it exists.

The threshold level for the patient satisfaction indicator is still under consideration. It will be set based on the survey instrument selected and available data on patient satisfaction outcomes for PGPs and Medicare beneficiaries using that measure. We also considered setting minimum threshold target levels for all of the indicators which would also be required for PGPs to achieve in order to earn a bonus. For example, requiring at least 35 percent of diabetics to have received a hemoglobin A1c test for a PGP to receive a bonus payment for that indicator. This would enable the demonstration to avoid paying bonuses to PGPs that can be considered low quality providers, at least as measured by these indicators. However, we believe that minimum target levels should not be required for this demonstration for two reasons. First, substantial numbers of beneficiaries will experience health care improvements if low performing groups increase

quality in response to the financial incentives provided by this demonstration. In other words, the PGP demonstration will provide broader benefits to the Medicare population by raising quality for low performing PGPs than by rewarding high performing PGPs for quality already achieved. Second, we are recommending larger improvements be required to achieve the quality improvement targets for PGPs starting at lower levels of quality, as discussed next.

Substantial Improvement-Over-Time Targets

- For nine of the eleven recommended indicators (all except the ACSC and patient satisfaction indicators), quality improvement over time will be demonstrated if the PGP's performance improves by 10 percent of its quality "deficit" from the previous year. The quality deficit is defined as the ideal rate (100%) minus the actual rate.

For example, if 50 percent of a PGP's diabetics had hemoglobin A1c tests in one year, it would have to raise that level to 55 percent the following year to demonstrate it had met the quality improvement target for that indicator. (A 50 percent deficit means 5 percentage points improvement required.) For these nine indicators, the ideal rate should be 100 percent since they are recommended for all Medicare beneficiaries with the indicated condition (absent identified exclusions).

An advantage of this approach, which is also used by Medicare's Quality Assessment and Performance Improvement (QAPI) program, is that it requires smaller improvements from PGPs that are already performing relatively well on an indicator and thus will have a harder time improving significantly. (For example, a PGP at 70 percent performance on hemoglobin A1c tests for diabetics would only need to improve to 73 percent in the following year.)

One limitation on this approach is that it is difficult to apply for the three indicators which are measured over two-year periods (e.g., mammograms). For those indicators four years of data are required since a two-year performance period must be compared with a previous two-year base period to demonstrate improvement over time. As a result, to limit the data analysis required for the PGP demonstration, we recommend that the substantial improvement over time target only be applied in the third year of the demonstration for the quality indicators measured over two-year periods. In the third year the required four years of claims data will be available (three performance years plus one base year). The fixed threshold targets will still apply in all three years of the demonstration for those indicators.

For the ACSC indicator an ideal rate has not yet been determined. As a result:

- A 10 percent reduction over the previous year's ACSC admission rate will qualify as meeting the quality improvement target.

For example, if a PGP had a rate of 40 ACSC admissions per 1,000 Medicare beneficiaries in a base year, then it will be required to reduce that rate to 36 admissions per 1,000 in the performance year to meet this target. The 10 percent criterion may be adjusted, however, depending on the statistical properties of ACSCs found in a PGP demonstration simulation we will conduct later this year. We will evaluate the likelihood of a 10 percent reduction occurring by chance, and recommend a larger percentage reduction target if this appears likely to happen.

Age and gender have been shown to affect the rates of ACSC admissions per 1,000 Medicare beneficiaries (HER unpublished data). Thus the rates of ACSC admissions per 1,000 beneficiaries for the base year will be age and gender adjusted to the characteristics of the performance year before the percentage improvement is calculated. To conduct the age and gender adjustments we will use the direct standardization method, which has been widely applied to disease and service utilization rates in health care (Klein & Schoenborn, 2001; Trisolini *et al.*, 2001).

Table 18-1 summarizes the threshold and improvement targets for each of the recommended quality indicators.

Recommendation 19: The Portion of the PGP Bonus Pool That Can be Earned Through Quality Performance

- HER recommends that 30 percent of the PGP Bonus Pool be reserved as the portion that can be earned through quality performance.

For the PGP demonstration, it is assumed that at least half of the PGP Bonus Pool will be reserved for performance incentives resulting from cost savings. However, that still leaves a fairly wide range of possibilities (0 percent to 50 percent) for the quality portion of the PGP Bonus Pool. There are three reasons for recommending 30 percent of the PGP Bonus Pool as the amount reserved for quality improvements. First, there is a mandate to include incentives for quality improvement in the BIPA 2000 legislation, so a significant percentage is warranted. Moreover, the IOM recently called for pilot testing

Table 18 - 1

Recommended Quality Performance Targets For Each Quality Indicator

Quality Indicator	Substantial Improvement Over Time Target	Pre-determined Threshold Target
Eye exams every 2 years for diabetics	10 percent improvement over the deficit from 100 percent compliance*	75 percent compliance
Hemoglobin A1c test every year for diabetics	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Lipid profile test every 2 years for diabetics	10 percent improvement over the deficit from 100 percent compliance*	75 percent compliance
Mammogram every 2 years for women aged 52-69	10 percent improvement over the deficit from 100 percent compliance*	75 percent compliance
Chest radiograph and electrocardiogram ≤ 3 months after initial diagnosis of CHF	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Left ventricular ejection fraction test during the current year for beneficiaries hospitalized with a principal diagnosis of CHF during the current year.	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Physician visit every 6 months for beneficiaries with chronic stable angina, COPD, CHF, or diabetes	10 percent improvement over the deficit from 100 percent compliance	90 percent compliance
Annual influenza vaccinations for all beneficiaries age 65 or older	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Pneumonia vaccination status for all beneficiaries age 65 or older	10 percent improvement over the deficit from 100 percent compliance	75 percent compliance
Rate of ACSC admissions per 1000 Medicare beneficiaries	10 percent reduction from the previous year's rate	National average rate for FFS beneficiaries
Patient satisfaction (specific indicator to be determined)	N/A	To be determined

*For the quality indicators measured over a 2 year period, the substantial improvement over time target will only be available during the third year of the demonstration. At that point, four years of data will be available (3 performance years plus 1 base year) to enable improvement calculations to be performed.

of public sector payment mechanisms with explicit quality improvement incentives for health care providers (IOM, 2001a). It noted that very little attention has been paid to development of ways to align payment incentives with quality improvement. The PGP demonstration represents an opportunity to demonstrate how cost saving and quality improvement incentives can be jointly applied in public sector payment mechanisms. Over time, the percentage allocated to quality performance could rise as PGPs become more familiar with these new types of incentives.

Second, the recommended level of 30 percent should provide sufficient incentives to motivate PGPs to take action to improve quality. If 30 percent of the PGP Bonus Pool is reserved for quality performance incentives, that means that up to 24 percent (30% quality share x 80% Sharing Rate) of Medicare Savings can be earned through quality performance. (Assuming 20% of Medicare Savings goes to CMS and the other 80% constitutes the PGP Bonus Pool.) HER simulations have shown that this provides for significant bonus payments to be earned by PGPs. At the same time, this approach reserves the majority (70%) of the PGP Bonus Pool for the efficiency incentives that are the primary focus of the demonstration.

Third, Medicare is protected financially since the actual payments made to PGPs based on quality performance will depend on the total amount of Medicare Savings they have achieved. Quality improvement bonuses will not be paid by CMS if Medicare Savings are not generated by PGPs to fund them.

One possible disadvantage of providing significant direct financial incentives for quality improvement is that they are relatively unexplored territory for public sector

payment systems. As a result, most PGPs will be unfamiliar with them in this context. This concern is mitigated by the overall design of the demonstration, however, since PGPs can only benefit from its incentives. They will not be subject to financial penalties if their performance does not reach their quality improvement targets.

Moreover, PGPs may welcome financial incentives, which support their efforts toward quality improvement. As the IOM (2001a) has stressed, existing payment policies often work against efforts of physicians to improve quality (including PGPs paid through FFS reimbursement). The IOM provided an interesting anecdote to illustrate this point, which is relevant to the PGP demonstration:

A physician group paid primarily on a fee-for-service basis instituted a new program to improve blood sugar control for diabetic patients. Specifically, pilot studies suggested that tighter diabetic management could decrease hemoglobin A1c levels by 2 percentage points for about 40 percent of all diabetic patients managed by the physician group. Data from two randomized controlled trials demonstrated that better sugar controls should translate into lower rates of retinopathy, nephropathy, peripheral neurological damage, and heart disease. The savings in direct health care costs (i.e., reduced visits and hospital episodes) from avoided complications have been estimated to generate a net savings of about \$2,000 per patient per year, on average, for 15 years. Across the more that 13,000 diabetic patients managed by the physician group, the project had the potential to generate over \$10 million in net savings each year. The project was costly to the medical group in two ways. First, expenses to conduct the project, including extra clinical time for tighter management, fell to the physician group. Second, over time, as diabetic complications fell, the project would reduce patient visits and, thus, revenues as well. But the savings from avoided complications would accrue to the insurer or a self-funded purchaser. (IOM, 2001a: 17)

If the percentage of the PGP Bonus Pool allocated to quality incentives were low or zero, then some of the quality improvement disincentives illustrated in the above anecdote would still be mitigated by the PGP demonstration payment methods, but many

would not. A bonus pool awarded solely for cost savings performance would reduce or prevent the losses the PGP would suffer due to reduced patient visits in the medium term, but would not reward the group for the extra time spent by physicians and others on quality improvement efforts in the short run, before utilization reduction effects were realized. Moreover, due to the risk of plan or physician group switching by beneficiaries, a PGP might not realize longer-term cost-efficiency bonuses from the PGP Bonus Pool due to the reduced utilization resulting from lower complication rates. Indeed, even fully capitated groups often cannot count on long-run cost savings from preventive care interventions due to the problem of patient switching. As a result, capitated groups often avoid investments in quality improvement or preventive care for fear that they will not realize financial returns on their investments, or, even worse, that their competitors may actually benefit if beneficiaries switch to the competitors' PGPs or health plans after the initial investments in preventive care have been made. Providing direct financial incentives for quality improvement in this demonstration will enable PGPs to reap financial returns more reliably, in the short-run, for investments required to improve quality indicators.

Recommendation 20: Calculating and Allocating Process and Outcome Quality Improvement Bonuses

- Participating PGPs will earn 1/8th of the quality portion of the PGP Bonus Pool for achieving a target for each of the eight final quality indicators selected for the demonstration.

Our recommended approach enables a PGP to earn its quality improvement bonus in the most flexible way, in eight discrete segments of 1/8th each. Thus achieving a quality target is accorded equal weight for each indicator. If targets are met for all eight

indicators, the PGP will earn the full 30 percent of the PGP Bonus Pool available for quality incentives. If targets are only met for some of the eight indicators, then the percentage of the PGP Bonus Pool earned by the PGP will be reduced proportionately. However, some portion of the PGP Bonus Pool will be earned for each target achieved, so PGPs will have continuing incentives for quality improvement, even if they believe they are not able to meet the targets for some of the indicators. A 1/8th portion represents 12.5% of the quality portion of the PGP Bonus Pool, 3.75% of the PGP Bonus Pool (12.5% share of quality portion of PGP Bonus Pool x 30% quality share of PGP Bonus Pool), and 3% of Medicare Savings (12.5% share of quality portion of PGP Bonus Pool x 30% quality share of PGP Bonus Pool x 80% PGP Bonus Pool share of Medicare Savings).

Although a higher weight might be assigned for some of the quality targets (e.g., patient satisfaction), we recommend for equal weights for two reasons. First, equal weighting is simple and avoids further complexity for implementing the demonstration. Second, it is not obvious how to differentially weight the quality indicators. Several factors could be considered for assigning weights. Weights could be based on equating the likelihood of achieving a pre-determined rate of quality improvement for each indicator. Those probabilities would be derived from the standard deviations of the indicators. Alternatively, the underlying utility gains that patients assign to pre-determined quality improvements could be assessed, such as the value of raising patient satisfaction 5 percent versus the value of a 5 percent reduction in the risk of blindness through providing eye exams to diabetics. We believe that both approaches are promising

and that work should be actively pursued to develop methodologies for assigning weights to different quality indicators. However, in the absence of widely-accepted weighting methodologies at the present time, we recommend using equal weights for this demonstration.

References

Adamache K, C. Liu and G Pope: *Simulations of Selected GVPS Design Parameters Using Historical Data*. Prepared for the U.S. Health Care Financing Administration. Waltham, Massachusetts: Health Economics Research, Inc., 1998.

AHRQ. *Provider Group CAHPS[®] Project Summary*. Prepared for the Group CAHPS National Stakeholder Meeting. Rockville, Maryland: U.S. Agency for Healthcare Research and Quality, 2001.

AHRQ. *CAHPS[®] 2.0 Survey and Reporting Kit*. Publication No. 99-0039A. Rockville, Maryland: U.S. Agency for Healthcare Research and Quality, 1999.

AMGA. *AMGA Patient Satisfaction Benchmarking Program*. Alexandria, Virginia: American Medical Group Association.
www.amga.org/AMGA2000/QMR/PSAT/index_psat.htm. Accessed January 9, 2002.

Asch S, E. Sloss, C. Hogan, *et al.*: Measuring Underuse of Necessary Care Among Elderly Medicare Beneficiaries Using Inpatient and Outpatient Claims. *Journal of the American Medical Association* 284(18): 2325-2333, 2000.

Centers for Medicare & Medicaid Services: (2001a) *HCQIP – National Clinical Topics. Diabetes Project Description*. <http://www.hcfa.gov/quality/11a3-c.htm>. Accessed: November 6, 2001.

Centers for Medicare & Medicaid Services: (2001b) *HCQIP – National Clinical Topics. Breast Cancer Project Description*. <http://www.hcfa.gov/quality/11a2-c.htm>. Accessed: November 6, 2001.

Centers for Medicare & Medicaid Services. Payments to Medicare+Choice (M+C) organizations. Chapter 7 in: *Medicare Managed Care Manual*. Washington: USDHHS, 2001. Available at: http://www.hcfa.gov/pubforms/86_mmc/mc86c07exhibit6.htm. [Accessed April 13, 2002].

Chassin M: Improving the Quality of Care. *New England Journal of Medicine* 335: 1060-1063, 1996.

Chassin M, R Galvin, *et al.*: The Urgent Need to Improve Health Care Quality. *Journal of the American Medical Association* 280(11): 1000-1005, 1998.

CMS. *Health Care Quality Improvement Program: Medicare Priorities*. Baltimore: U.S. Centers for Medicare & Medicaid Services, 2000.

Current Procedural Terminology CPT 2002, American Medical Association, AMA Press, 2001.

Davies S, J Geppert, M McClellan, *et al.*: *Refinement of the HCUP Quality Indicators*. AHRQ Publication No. 01-0035. Rockville, Maryland, 2001.

Eddy D: Performance measurement: problems and solutions. *Health Affairs* 17(4) 7-25, 1998.

Ellis R, G Pope, *et al.*: 1996, Diagnosis-Based Risk Adjustment for Medicare Capitation Payments. Health Care Financing Review, Spring.

Fraser I, P McNamara, G Lehman, *et al.*: The Pursuit of Quality by Business Coalitions: A National Survey. *Health Affairs* 18(6): 158-165, 1999.

Garnick D, J Fowles, A Lawthers, *et al.*: Focus on quality: profiling physicians' practice patterns. *Journal of Ambulatory Care Management* 17(3): 44-75, 1994.

Garnick DW, HS Luft, JC Robinson, J Tetreault: Appropriate Measures of Hospital Market Areas, *Health Services Research* 22:1, April 1987.

Goldstein E., Cleary P., Langwell K., *et al.*: Medicare Managed Care CAHPS[®]: A Tool for Performance Improvement. *Health Care Financing Review* 22(3): 101-107, 2001.

Institute of Medicine: *Crossing the Quality Chasm*. Washington: National Academy Press, 2001a.

Institute of Medicine: *Envisioning the National Health Care Quality Report*. Washington: National Academy Press, 2001b.

Integrated Healthcare Association. *Pay for Performance*. www.iha.org [Accessed January 23, 2002.] Walnut Creek, California, 2002.

Jencks S, T Cuerdon, D Burwen, *et al.*: Quality of Medical Care Delivered to Medicare Beneficiaries: A Profile at State and National Levels. *Journal of the American Medical Association* 284(13): 1670-1676, 2000.

Johantgen M, A Elixhauser, J Ball, *et al.*: Quality Indicators Using Hospital Discharge Data: State and National Applications. *The Joint Commission Journal on Quality Improvement* 24(2): 88-105, 1998.

Klein R and Schoenborn C: *Age Adjustment Using the 2000 Projected U.S. Population*. Statistical Notes No. 20. Atlanta: Centers for Disease Control, 2001.

Konstvedt P: *The Managed Health Care Handbook, Fourth Edition*. Gaithersburg, Maryland: Aspen Publishers, Inc., 2001.

KPMG Consulting: *Physician Group Practice Demonstration Selection Tool*. Contract No. 500-97-0441, Task Order No. 2014. Prepared for the Centers for Medicare & Medicaid Services, 2001a.

KPMG Consulting: *Value Based Purchasing Project: Recommendations for Conference Speakers*. Contract No. 500-97-0441, Task Order No. 2014. Prepared for the Centers for Medicare & Medicaid Services, 2001b.

McCall N, J Harlow and D Dayhoff: Rates of Hospitalization for Ambulatory Care Sensitive Conditions in the Medicare+Choice Population. *Health Care Financing Review* 22(3): 127-145, 2001.

McCall, Pope, *et al.*, *Research and Analytic Support for Implementing Performance Measurement in Medicare Fee for Service*, CMS Contract No. 500-95-0058, Second Annual Report, January 6, 2000.

McCall N, G Pope, K Adamache, *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 Master Order Contract No. 500-95-0058, Task Order No. 2. Waltham, Massachusetts: Health Economics Research, Inc., 1998.

NCQA. HEDIS 3.0. Washington: *National Committee for Quality Assurance*, 1997.

NCQA. HEDIS 2002. Washington: *National Committee for Quality Assurance*, 2001.

Newcomer L: Physician, Measure Thyself. *Health Affairs* 17(4) 32-35, 1998.

Parchman M and S Culler: Preventable Hospitalizations in Primary Care Shortage Areas: An Analysis of Vulnerable Medicare Beneficiaries. *Archives of Family Medicine* 8: 488-491, 1999.

Pope G, B Wu and N McCall: *Refinement and Validation of Medicare+Choice Plan Ambulatory Care Sensitive Condition Admission Rates*. Final Report Prepared for the Centers for Medicare & Medicaid Services, Contract No. CMS-500-95-0058, Task A. Waltham, Massachusetts: Health Economics Research, Inc, 2002.

Pope G and J Chromy: *Issues in Demonstrating Group-Specific Volume Performance Standards*. Prepared for the U.S. Health Care Financing Administration. Waltham, Massachusetts: Health Economics Research, Inc., 1997.

Pope G *et al.*: *Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment*. Final Report to HCFA. Waltham, MA: Health Economics Research, Inc., 2000.

Sanderson-Austin J and H Wetzler: Room for Improvement: Data from AMGA's Patient Satisfaction Survey Reveal Variation by Specialty. *Group Practice Journal*, January, 2001.

Sennett C: Moving Ahead, Measure by Measure. *Health Affairs* 17(4) 37-38, 1998.

Shortell S, J Zazzali, L Burns, *et al.*: Implementing Evidence-Based Medicine: The Role of Market Pressures, Compensation Incentives, and Culture in Physician Organizations. *Medical Care* 39(7): I-62 – I-78, 2001.

Tompkins C, S Bhalotra, M Glavin, *et al.*: *A Demonstration Of Group-Specific Volume Performance Standards (GVPS): Design Report*. Waltham, Massachusetts: Institute for Health Policy, Brandeis University, 1996a.

Tompkins C, S Wallack, S Bhalotra, *et al.*: Bringing Managed Care Incentives to Medicare's Fee-For-Service Sector. *Health Care Financing Review* 17(4): 43-63, 1996b.

Trisolini M, Klosterman M, McCall N, *et al.*: *Medicare Quality Monitoring System Version 1 Specifications: M4 – Frequency, Rate, Length of Stay, and Costs of Hospital Discharges for HCQIP Clinical Priority Areas of Heart Failure, Stroke, Pneumonia, and AMI by Age Group, Sex, Race, Medicare Eligibility Status, Medicaid Enrollment Status, and Geographic Area, Calendar Year 1999*. Prepare for the Centers for Medicare & Medicaid Services. CMS Contract No. 500-95-0058, T.O. No. 16. Waltham, Massachusetts, Health Economics Research, Inc., 2001.

Weiner J, S Parente, D Garnich, *et al.*: Variation in Office-Based Quality: A Claims-Based Profile of Care Provided to Medicare Patients With Diabetes. *Journal of the American Medical Association* 273(19): 1503-1508, 1995.

Wu B. and G. Pope. Left ventricular ejection fraction test rates for Medicare beneficiaries with heart failure. *American Journal of Medical Quality* 17(2): 61-66, 2002.

Zaslavsky A, Zaborski L, Ding L, *et al.*: Adjusting Performance Measures to Ensure Equitable Plan Comparisons. *Health Care Financing Review* 22(3): 109-126, 2001.

Zwanziger J, G Melnick, KM Eyre: Hospitals and Antitrust: Defining Markets, Setting Standards, *Journal of Health Politics, Policy and Law* 19(2), Summer, 1994.

Zwanziger J, G Melnick, JM Mann: Measures of Hospital Market Structure: A Review of the Alternatives and a Proposed Approach, *The Journal of Socioeconomics Planning Sciences*, 24(2):81-95, 1990.

Appendix A

Patient Assignment Steps

Appendix A

Patient Assignment Steps

As discussed in *Recommendation 2: Assigning Beneficiaries to PGPs*, the beneficiary assignment criteria requires that a PGP provide at least one E&M service to a beneficiary, and more E&M services (as measured by Medicare payments) than any other physician practice (group or solo). The patient assignment steps for a given PGP are outlined below. These steps will be followed for both base and performance years. In Step 1, the PGP is required to provide its Provider Tax Number(s) so that its Medicare claims can be identified. In Step 2, all PGP claims containing at least one evaluation and management (E&M) service are identified.³⁵ A beneficiary with a claim in this set meets the first condition of the assignment criteria, i.e., the PGP provides at least one E&M service to the beneficiary. Beneficiaries with any Medicare+Choice enrollment during the year are not eligible for assignment. The second condition of the assignment criteria requires that the PGP provide more E&M services to the beneficiary than any other provider. To ascertain whether the second condition is met, Step 2 also calculates the Medicare payments for all qualifying E&M services provided by the PGP to the beneficiary. Then in Step 3, for physician practices other than the PGP, claims containing at least one E&M service are identified. For each physician practice, the Medicare payments for all E&M services provided to the beneficiary are calculated. Finally, in Step 4 the beneficiary is assigned to the PGP if the Medicare payments for E&M services provided by the PGP is greater than that for any other physician practice.

Step 1

The PGP demonstration participant will provide to CMS its Provider Tax Number(s).

Step 2

Pull all physician/supplier Part B claims that contain both the Provider Tax Number provided in Step 1 and at least one E&M procedure code.

Determine the set of unique HICNOs for the set of claims. Eliminate claims corresponding to beneficiaries with any Medicare+Choice enrollment. For each HICNO, determine the Medicare payments for all qualifying E&M services.

Step 3

For the HICNOs in Step 2, pull all physician/supplier Part B claims. Not including the Provider Tax Number(s) provided in Step 1, determine the set of unique Provider Tax

³⁵ As discussed in *Recommendation 2: Assigning Beneficiaries to PGPs*, certain E&M services such as consultations, emergency department visits, and critical care services do not reflect the PGP's ability to manage the health care of a beneficiary, and are thus not used in the assignment.

Numbers for the set of claims. For each Provider Tax Number and HICNO, determine the Medicare payments for all qualifying E&M services.

Step 4

Assign HICNOs in Step 2 to the PGP if for each Provider Tax Number in Step 3, the Medicare payments for all qualifying E&M services in Step 2 is greater than that in Step 3.

Appendix B

Steps in Defining the Comparison Population and Calculating the Target Growth Rate

Appendix B

Steps in Defining the Comparison Population and Calculating the Target Growth Rate

We begin by assuming that the beneficiaries assigned to a participating PGP have been identified. We call these beneficiaries "assigned beneficiaries". This appendix describes the steps in defining the comparison populations for this PGP. These steps will be followed for both base and performance years to define comparison populations for both years. The comparison population need not consist of the same individuals in both base and performance years.

Step 1: Match HICNOs of assigned beneficiaries against denominator file to identify SSA county code of residence.

Step 2: Identify counties in the PGP's market area, i.e., counties where at least 1% of PGP beneficiaries reside.

Step 3: Access denominator file to identify all Medicare beneficiaries residing in these counties.

Step 4: From the beneficiaries identified in Step 3, eliminate those with any managed care (Medicare + Choice) enrollment during the year, and eliminate those assigned to the participating PGP. The remaining beneficiaries are the "preliminary comparison group".

Step 5: Verify that the preliminary comparison group contains an effective sample size³⁶ of at least 30,000 beneficiaries and is not skewed or unrepresentative of the local Medicare population. If so, proceed to Step 6. If not, return to Step 2 and add more counties to the PGP's market area in descending order of number of assigned PGP beneficiaries until the preliminary comparison group contains an effective sample size of at least 30,000 beneficiaries. Determining effective sample size is described in Appendix F.

Rationale: We expect that approximately 30% of the preliminary comparison group will not have any qualifying evaluation and management claims in a one-year period. We will eliminate such "non-users" in Step 6 to obtain the "final comparison group", which we want to contain an effective sample size of at least 15,000 beneficiaries. To ensure that the final comparison group will contain an effective sample size of at least 15,000 beneficiaries, we require that the preliminary comparison group contain an effective sample size of at least 30,000 beneficiaries ($15,000/0.70 = 21,429$, and we have rounded

³⁶ The effective sample size differs from the nominal sample size in accordance with the weighting of counties by PGP beneficiary proportions. As more weight is put on counties with smaller populations, the effective sample size declines. See Appendix F.

up to 30,000 to provide some margin for error in the proportion with an evaluation and management claim).

Step 6: Pull off 100% of Medicare FFS claims for comparison group beneficiaries identified in Step 4 (DESY run.) Eliminate beneficiaries with no qualifying evaluation and management claims ("nonusers") from the comparison group. The remaining beneficiaries are the "final comparison group". Verify that the final comparison group contains an effective sample size of at least 15,000 beneficiaries in both base and performance years.

Step 7: Verify that no major changes in the comparison group have occurred from the base to the performance year (e.g., because of a large change in Medicare managed care penetration in the participating PGP's market area).

Step 8: Compute base and performance year per capita expenditures for each county in the market area using the claims obtained in Step 6.

Step 9: Compute a county risk adjustment factor for each county in the market area for both base year and performance year. Adjust base year county per capita expenditures for performance year health status casemix by multiplying base year expenditures by the ratio of the county risk adjustment factor in the performance year to the county risk adjustment factor in the base year. This adjustment will be done for each county in the market area, county by county.

Step 10: Calculate base and performance year averages of county per capita expenditures weighting by proportions of PGP beneficiaries in each market area county in the base and the performance years, respectively. Use risk-adjusted base year county per capita expenditures from Step 9 in this calculation. Use unadjusted performance year county per capita expenditures in this calculation. Then compute the growth rate in weighted average per capita expenditures from the base to the performance year. The result of this step is the target expenditure growth rate.

Time estimates after data access permissions have been granted, best-case scenario:

Steps 1-4: one month.

Steps 5-6: one month.

Steps 7-10: one month.

Total: three months.

Appendix C

Potential Medicare Claims-Based Process and Outcome Quality Indicators

Appendix C

Potential Medicare Claims-Based Process and Outcome Quality Indicators

PROCESS INDICATORS

Necessary Care

One visit every 6 months for patients with chronic stable angina
One visit every 6 months for patients with chronic obstructive pulmonary disease
Chest radiograph <3 months after initial diagnosis of congestive heart failure
Electrocardiogram <3 months after initial diagnosis of congestive heart failure
One visit every 6 months for patients with congestive heart failure
Frequent serum potassium test for patients with congestive heart failure
One visit every 6 months for patients with ischemic heart disease
Glycosylated hemoglobin or fructosamine test every 6 months for patients with diabetes
Urinalysis every year for patients with diabetes
Triglycerides test every year for patients with diabetes
Total cholesterol test every year for patients with diabetes
HDL cholesterol test every year for patients with diabetes
Eye examination every year for patients with diabetes
Influenza vaccine every year for patients with diabetes
Foot exam every year for patients with diabetes
One visit every six months for patients with diabetes
Two visits per year for patients with hypertension
One visit per year for patients with osteoarthritis

Preventive Care

One visit every year for all Medicare beneficiaries
Assessment of visual impairment at least once every 2 years for all Medicare beneficiaries
Mammography every two years for all female Medicare beneficiaries aged 50-75

Appendix C (continued)

Potential Medicare Claims-Based Process and Outcome Quality Indicators

OUTCOME INDICATORS

Avoidable Morbidity

Among patients with known COPD, subsequent admission for a respiratory diagnosis

Nonelective admission for congestive heart failure

Among patients with known diabetes, admission for hyperosmolar or ketotic coma

Surgery for a perforated appendix

Admissions for other types of “ambulatory care sensitive conditions” (ACSCs)

Immunization-preventable pneumonia or influenza

Among patients with pneumonia, diagnosis of lung abscess or empyema

Among patients with known angina, >2 ED visits for cardiovascular-related diagnoses in 1 year

Surgical site infection within 30 days (for selected surgical procedures)

Readmission rates at 2 days, 7 days, and/or 30 days following major surgery

Avoidable Mortality

In-hospital mortality following hysterectomy, laminectomy, cholecystectomy, TURP, hip replacement, or knee replacement

30-day mortality following acute AMI

30-day mortality following elective outpatient cardiac stress test

30-day mortality following elective outpatient colonoscopy

30-day mortality following elective outpatient surgical procedures

SOURCE: Asch *et al.*, 2000; Jencks *et al.*, 2000; Johantgen *et al.*, 1998; McCall *et al.*, 1998; McCall *et al.*, 2001; Garnick *et al.*, 1994; Institute of Medicine 2001b

Appendix D

Ambulatory Care Sensitive Conditions

Appendix D

Ambulatory Care Sensitive Conditions

Asthma/COPD
Congestive Heart Failure
Seizure Disorder
Diabetes Mellitus
Hypertension
Gastric or Duodenal Ulcer
Hypoglycemia
Urinary Tract Infections
Cellulitis
Dehydration
Hypoalkemia
Pneumonia
Severe Ear/Nose/Throat Infections
Influenza
Malnutrition

SOURCE: McCall *et al.*, 2001

Appendix E

Simulations of Adjusted Sharing Rate for Teaching Hospitals; Discussion of Timeliness of IME Payments

Appendix E

1) Simulations of Adjusted Sharing Rate for Teaching Hospitals;

2) Discussion of Timeliness of IME Payments

I. Simulations of Adjusted Sharing Rate for Teaching Hospitals

A set of simulations were performed using the older GVPS methodology (Tompkins *et al.*, 1996) rules and methods for handling foregone IME and DSH payments due to fewer admissions at a teaching hospital that is a partner with a PGP in the demonstration (see text discussion Recommendation 11). The objective of the simulations was to determine whether the older GVPS methodology over-, under-, or just compensates participating PGPs for foregone IME and DSH payments. As described in text Recommendation 11, that method involves adjusting the sharing rate by adding to the base sharing rate the IME and DSH share of total Medicare per capita expenditures (PCE).

The IME and DSH add-on share, hereafter *add-on PCE share*, is the ratio on the left-hand side of Equation 11-1 below. It is the share of IME and DSH in total per capita expenditures. As shown in Equation 11-1, the *add-on PCE share* can be decomposed into the product of the PPS hospital share of total expenditures per patient (*hospital PCE share*) and IME and DSH share of total PPS hospital payments (*add-on hospital share*).

$$\underbrace{\frac{\text{PPS Add - On Payments}}{\text{Total Expenditures for Patients}}}_{\text{"Add-On PCE Share"}} = \underbrace{\frac{\text{PPS Hospital Payments}}{\text{Total Expenditures for Patients}}}_{\text{"Hospital PCE Share"}} \times \underbrace{\frac{\text{PPS Add - On Payments}}{\text{PPS Hospital Payments}}}_{\text{"Add-On Hospital Share"}} \quad (11-1)$$

Both the hospital PCE share and the add-on hospital shares can vary by potential participants, so they were both varied to obtain a range of add-on PCE shares.

Using PPS FY 2002 Impact File data, a range of teaching hospital add-on shares and situations were estimated: (1) a specific large teaching hospital in a large urban setting that is part of a multi-partner delivery system, (2) a “typical” teaching hospital located in a large urban area, and (3) a “typical” teaching hospital located in an “other” urban area. The estimated add-on hospital shares were, respectively, 31.8 percent, 19.3 percent, and 11.8 percent. A hospital PCE share of 55 percent was used, resulting in three add-on PCE shares: 0.06 for a typical teaching hospital in an “other” urban area, 0.11 for a typical teaching hospital in a large urban area, and 0.19 for a participating PGP teaching hospital. In other words, PPS IME and DSH payments are estimated to comprise 6%, 11%, and 19% of total per capita expenditures for beneficiaries assigned to a participating PGP in these three scenarios. These shares were added to a base sharing rate of 0.80 to obtain three adjusted sharing rates: 0.86, 0.91, and 0.99. Because the last sharing rate exceeded the M+C program’s 0.95 sharing rate, it was capped at 0.95.

Medicare savings per patient were estimated for four Medicare savings rates (0.03, 0.05, 0.07, and 0.10). Total bonuses per patient were then calculated for the four Medicare savings rates and the three adjusted sharing rates. For a given Medicare sharing rate, the marginal (incremental) bonus was calculated by subtracting the total bonus based on the base 0.80 sharing rate (no add-on adjustment) from the total bonus for a sharing rate adjusted for the IME/DSH add-on percentage. For instance, at a seven percent Medicare savings rate, the total simulated bonus per assigned beneficiary at the 0.80 sharing rate is \$309.40 and the total bonus per beneficiary at the 0.91 sharing rate is \$350.52 (the typical teaching hospital in a large urban area). The marginal bonus is then

\$41.12 per patient. Twelve marginal bonuses per patient by Medicare savings rate and the adjusted sharing rate are shown in the top panel of Table E-1.

The actuarial estimates of the gross add-on losses per patient are shown in the middle panel of Table E-1. Losses were estimated using the following steps. Per beneficiary add-on payments were actuarially estimated as the product of the add-on PCE share (see Equation 11-1), PCE, and the participant's share of PCE. For this step, PCE was assumed to be \$6,500 per patient and the participant's share of PCE was assumed to be 40 percent. Next, actuarial estimates of the gross add-on losses due to participating in the demonstration were estimated as the product of the estimated per beneficiary add-on payments, the IME and DSH share of PPS hospital payments, and the Medicare savings rate. Losses were estimated for the same four Medicare savings rates previously used and the three IME and DSH shares.

The net add-on per capita payments are shown in the bottom panel of Table E-1. They were derived by subtracting the gross add-on losses from the marginal bonuses. The net add-on per capita payments are all positive, showing that the older GVPS approach to adjusting the sharing rate over-compensates PGPs for lost IME and DSH payments due to fewer admissions in teaching hospitals. Given the changes to the older GVPS methods for calculating bonus payments (e.g., elimination of the patient capture ratio), that methodology now only works in certain (unlikely) circumstances. The amount of the net add-on payment varies by the Medicare savings rate, the adjusted sharing rate, and other factors. To bring the net add-on payment back to zero would require many *ad hoc* adjustments to the sharing rates.

II. The Timeliness of IME Payments

The inclusion of IME in per capita expenditures, comparison growth rate, and the sharing rate could require post-demonstration-settlement IME values for not only PGP participants but teaching hospitals in the market area. The reason is that IME payments for a given payment period are based on Intern/Resident to Bed (IR-Bed) and Intern/Resident to Adjusted Daily Census (IR-ADC) ratios from a prior payment period. Payments based on such lagged values require settlement two to three years after the payment. The IME amounts included on the contemporaneous Medicare claims processed for the demonstration are only estimates. Time lags in settling hospital cost reports (2-3 years) and adjusting the claims data are too long to permit timely assessment of PGP performance.

For operating and capital IME payments, the foregoing implicitly assumes that the IR-Bed and IR-ADC ratios are subject to non-trivial annual variation. This implicit assumption might not hold any longer because of the Balanced Budget Act of 1997 (BBA). In response to long-term increases in the size of residency programs and, hence, Medicare graduate medical education payments, the BBA set individual hospital caps (ceilings) on the total number of interns and residents and on the IR-Bed and IR-ADC ratios to the values that each hospital had in 1996.

The BBA also established a voluntary transitional resident-reduction program for hospitals that wanted to decrease the size of their residency programs. The resident-reduction program would have made “hold harmless” payments to participating hospitals for up to five years as they reduced the size of their residency programs. HCFA published the specific qualifying regulations (*Federal Register*/Vol. 64, No. 159/August

18, 1999) with a one-time application deadline of November 1, 1999. Only two hospitals submitted applications and both at the time were in bankruptcy proceedings and facing permanent closure.

The weak response to the BBA's voluntary resident-reduction program suggests that most hospitals do not intend to deliberately downsize their residency programs in the near future. This history, thus, suggests that large residency programs will probably only experience small annual variations in the size of their residency programs. The only residency programs that might experience large annual changes in their size are likely to be small programs. And it is unlikely that hospitals with small residency programs are going to be that concerned with IME payments since the payments are a small part of their overall Medicare payments.

In sum, we do not believe that it is necessary for CMS to revise bonus computations under the PGP demonstration calculated with estimated IME payments on contemporaneous claims. If CMS were to do so, it would have to revise computations 2 to 3 years later with settled Medicare Cost Report IME values. We believe that this is neither necessary, desirable, nor feasible under the PGP demonstration.

Table E-1

**Simulated Effect on Per Capita Bonuses of Sharing Rate
Adjustments for PPS Add-Ons (IME/DSH)**

Medicare Savings Rate	Adjusted Sharing Rate¹		
	<u>0.86</u>	<u>0.91</u>	<u>0.95</u>
	I. Marginal Bonuses ²		
3%	\$10.77	\$17.62	\$24.86
5	17.94	29.37	41.44
7	25.12	41.12	58.01
10	35.89	58.74	82.88
	II. Add-on Gross Revenue Losses ³		
3	5.07	8.29	13.64
5	8.44	13.82	22.74
7	11.82	19.35	31.83
10	16.89	27.64	45.47
	III. Net Change in Bonus Per Capita ⁴		
3	5.70	9.33	11.22
5	9.50	15.55	18.70
7	13.30	21.77	26.18
10	19.00	31.10	37.40

NOTES:

¹Sharing rates based on, respectively, a typical hospital in "other" urban areas, a typical hospital in large urban areas, and an actual hospital that is part of an integrated delivery system. Sharing rate is increased beyond its base of 80% by the hospital's IME and DSH share.

²Equal to the total bonus for indicated sharing rate minus the total bonus based on the basic sharing rate of 0.80 (80%). The increase in bonus payments due to the upward adjustment in the sharing rate.

³Simulated lost revenue from foregone PPS add-ons due to lower hospital admissions (see text for explanation).

⁴Equal to marginal bonus minus the add-on gross revenue losses.

SOURCE: Health Economics Research, Inc.

Appendix F

Effective Sample Size of Market Area (Comparison Group) Beneficiaries

Appendix F

Effective Sample Size of Market Area (Comparison Group) Beneficiaries

The PGP demonstration target computation involves computing an expenditure growth rate from per capita expenditures of market area counties. For the target computation, the weights on the counties for the computation of the market area per capita expenditures will be each county's share of PGP beneficiaries (see Recommendation 3). In general, as shown below, these weights will lead to a larger variance of the estimated market area expenditure growth rate as compared to weighting by each county's share of total beneficiaries in market area counties. This may be seen intuitively in the case when the PGP share puts a large weight on a county with few beneficiaries, where the county per capita expenditure has a large variance. To understand the variance of the estimated growth rate weighting by the PGP shares, it is useful to measure the "effective sample size" of this estimate, in particular whether it exceeds the minimum sample size of 15,000 beneficiaries (Recommendation 7) desired for statistical stability of estimated expenditure growth rates. The effective sample size is the (smaller) sample size that gives the same variance when weighting by share of total beneficiaries per county as is attained when weighting by the share of PGP beneficiaries per county. This appendix analyzes the effective size for the computation of the market area growth rate.

The expenditure growth rate is computed as the weighted average of county per capita expenditures. For simplicity, in this analysis, we will assume only 2 counties.

(However, the analysis generalizes to any number of counties in a straightforward way.)

The two weights that we will consider for county expenditures are: (1) the market area share of total Medicare FFS beneficiaries in the county; and (2) the market area share of PGP beneficiaries in the county.

Let 1 denote county 1 with n_1 beneficiaries, 2 denote county 2 with n_2 beneficiaries, $n = n_1 + n_2 =$ total market area beneficiaries. X denotes individual beneficiary per capita expenditures, which has variance σ^2 . Mean expenditures in county 1 are denoted m_1 and m_2 denotes mean expenditures in county 2. $m_1 = \Sigma X/n_1$, $m_2 = \Sigma X/n_2$ where the summations are across beneficiaries in county 1 and in county 2, respectively.

We will consider the variance of the weighted sum of m_1 and m_2 :

weighted average county per capita expenditures = $E = w_1 m_1 + w_2 m_2$, where $w_1 + w_2 = 1$.

Note that when $w_1 = n_1/n$ and $w_2 = n_2/n$, then E is mean per capita expenditures for the n beneficiaries in the two counties. Now

$$\text{Var}(E) = (w_1)^2 \text{Var}(m_1) + (w_2)^2 \text{Var}(m_2) = (w_1)^2 \sigma^2/n_1 + (w_2)^2 \sigma^2/n_2. \quad (1)$$

Note that if $w_1 = n_1/n$ and $w_2 = n_2/n$, then

$$\text{Var}(E) = \sigma^2[n_1/n^2 + n_2/n^2] = \sigma^2/n. \quad (2)$$

To determine the effective sample size when the weights w_1 and w_2 do not equal n_1/n and n_2/n , respectively, we want to derive the sample size n that gives the same variance weighting by n_1/n and n_2/n as with arbitrary weights w_1 and w_2 . This is given by equating equations (1) and (2) and solving for n :

$$\sigma^2/n = (w_1)^2\sigma^2/n_1 + (w_2)^2\sigma^2/n_2.$$

The solution is

$$n_{\text{effective}} = 1/[(w_1)^2/n_1 + (w_2)^2/n_2]. \quad (3)$$

Note that if $w_1 = 1$ and $w_2 = 0$, then $n_{\text{effective}} = n_1$, and if $w_1 = 0$ and $w_2 = 1$, then $n_{\text{effective}} = n_2$, as expected. Also, if $w_1 = n_1/n$ and $w_2 = n_2/n$, then the above equation reduces to $n_{\text{effective}} = n$, as expected.

Example: Let's suppose $n_1 = 1,000$ and $n_2 = 2,000$. Then $n = 3,000$. The minimum variance weights (see below) are the share of beneficiaries in each county, i.e., $w_1 = n_1/n = 0.33$ and $w_2 = n_2/n = 0.67$. Instead, let's suppose that the share of PGP beneficiaries in the counties is such that we weight each county equally so that $w_1 = w_2 = 0.50$. Then $n_{\text{effective}} = 2,667$. Or suppose the PGP beneficiary shares are $w_1 = 0.90$ and $w_2 = 0.10$. Then $n_{\text{effective}} = 1,227$.

To compute the effective sample size for the market area, Equation (3) above should be used (generalized to more than 2 counties as necessary), where

w_i , $i = 1, \dots$, # counties in the market area, are the shares of PGP beneficiaries in each county in the market area;

n_i , $i = 1, \dots$, # counties in market area, are the number of comparison group beneficiaries residing in each county in the market area.

The goal is that the effective sample size of the market area is at least 15,000 so that statistically stable estimates of expenditure growth rates can be made.

Proof that weighting by the share of beneficiaries in each county minimizes the weighted variance

(differentiate the variance of E wrt w_1 and set to zero to solve for the w_1 that minimizes Var(E)):

$$\partial \text{Var}(E) / \partial w_1 = 2w_1\sigma^2/n_1 + 2(1-w_1)(-1)\sigma^2/n_2 = 0.$$

$$\Rightarrow w_1/n_1 - (1-w_1)/n_2 = 0.$$

$$\Rightarrow w_1 = (1/n_2) / [1/n_1 + 1/n_2].$$

$$\Rightarrow w_1 = 1/[n_2/n_1 + 1].$$

$$\Rightarrow w_1 = 1/[(n_2+n_1)/n_1].$$

$$\Rightarrow w_1 = n_1/(n_2+n_1).$$

$$\Rightarrow w_1 = n_1/n.$$