ANALYSIS OF 2011 PHYSICIAN FEEDBACK PROGRAM GROUP REPORTS

Wilfredo Lim, John Schurrer, and Mai Hubbard [*]

INTRODUCTION

In December 2012, the Centers for Medicare & Medicaid Services (CMS) produced and distributed confidential Physician Feedback reports to each of the 54 medical group practices that chose to participate in the 2011 Group Practice Reporting Option (GPRO) of the Physician Quality Reporting System (PQRS). Each report provided information on the quality of care and resource use for Medicare fee-for-service (FFS) beneficiaries treated by the medical groups in 2011. This summary provides descriptive statistics on the quality and cost of care displayed in those reports that the 54 groups provided to Medicare beneficiaries.

A. Description and Composition of the Group Practices Participating in the 2011 Group Practice Reporting Option

To participate in the GPRO of the 2011 PQRS program, a group practice had to be a single provider entity, identified by a single tax identification number (TIN). Additional participation criteria included the following:

- The group practice had at least 200 individual physicians or other medical professionals (identified by individual National Provider Identifiers, or NPIs) who had reassigned their billing rights to the TIN.
- The group practice submitted a self-nomination letter to CMS to participate in the 2011 GPRO PQRS program.
- CMS determined that the self-nominating group practice met the program definition of a group practice and complied with other program requirements.

CMS determined that 54 groups were eligible to participate in the 2011 PQRS GPRO, encompassing 37,745 eligible professionals. Eligible professionals include physicians and other medical professionals (such as physician assistants and nurse practitioners) who were Medicare-enrolled providers and who billed under the group practice's TIN in 2011.

On average, primary care physicians (PCPs) accounted for 22 percent of the group's affiliated eligible professionals. Medical specialists accounted for 22 percent; surgeons for 16 percent; emergency medicine physicians for 4 percent; other physicians for 13 percent; and other medical professionals for 23 percent of the group's affiliated eligible professionals. (A professional's medical specialty was determined based on the two-digit CMS medical specialty code listed most often on his or her 2011 Part B claims.) Although the average group practice profile was one in which PCPs were typically the plurality specialty, about one-fourth (26

[*] Mathematica Policy Research, 955 Massachusetts Avenue, Suite 801, Cambridge, MA 02139. Telephone: (617) 491-7900. Fax: (617) 491-8044. http://www.mathematica-mpr.com.

percent) of the group practices were composed of more medical specialists and surgeons than PCPs.

1. Attribution of Medicare Beneficiaries to Group Practices

For each of the 54 GPRO practices, CMS attributed a Medicare FFS beneficiary to the group if eligible professionals in the group billed for at least two of the beneficiary's eligible office visits or other outpatient evaluation and management (E&M) services provided in 2011 and the group practice billed the plurality of 2011 E&M charges for that beneficiary. Attribution-eligible E&M codes are listed in Table 1.

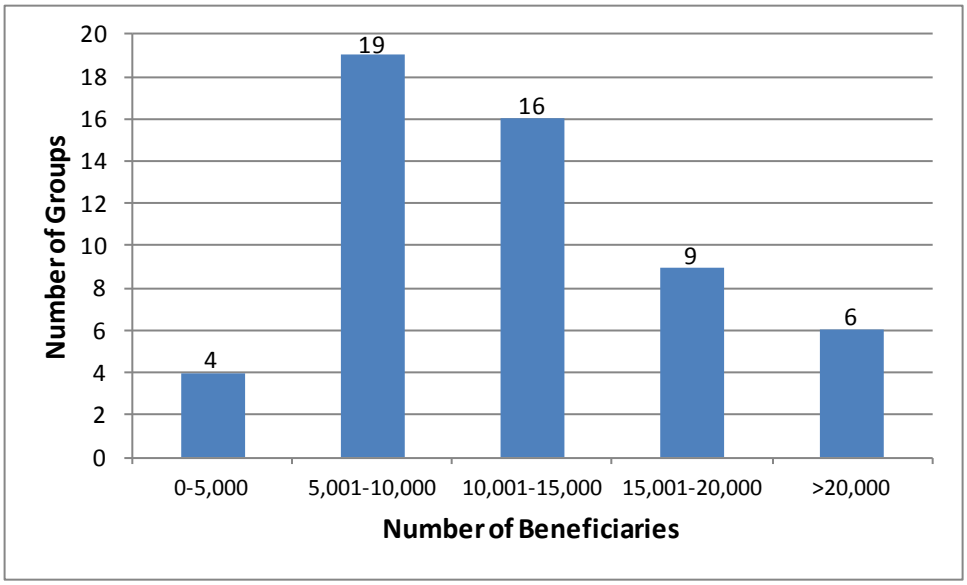
Table 1. Attribution-Eligible Medicare Part B Evaluation and Management Service Codes

Codes	Labels
99201	New patient, brief
99202	New patient, limited
99203	New patient, moderate
99204	New patient, comprehensive
99205	New patient, extensive
99211	Established patient, brief
99212	Established patient, limited
99213	Established patient, moderate
99214	Established patient, comprehensive
99215	Established patient, extensive

Note: The following E&M services were not considered when attributing beneficiaries to the group practices: hospital inpatient; nursing facility; care plan oversight; home care; domiciliary, rest home, or custodial care; consultations; emergency department; patient transport; critical care; neonatal intensive; newborn care; special evaluation and management; other E&M; preventive medicine; case management; prolonged; or hospital observations.

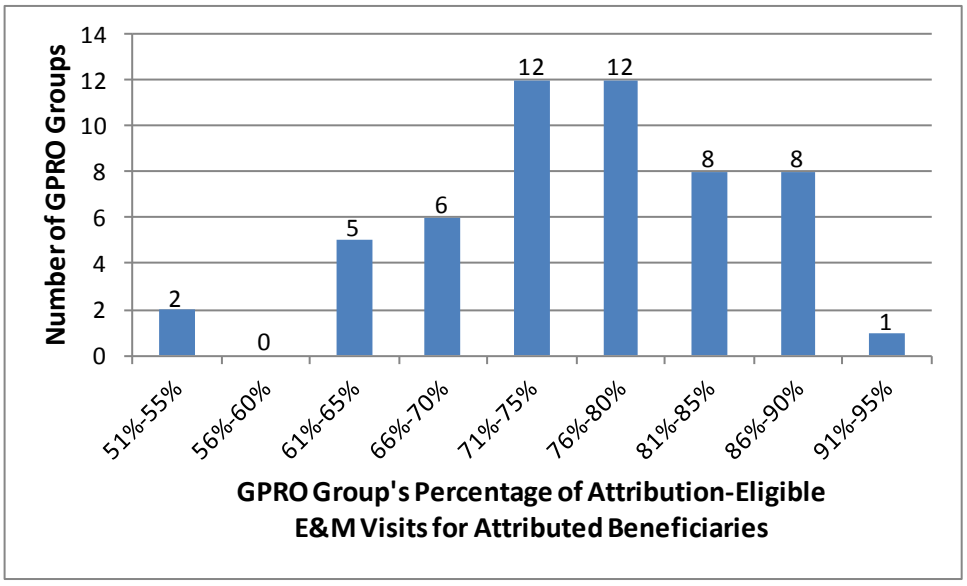
On average, 12,764 beneficiaries were attributed to groups, with 808 beneficiaries attributed to the smallest group practice and 33,907 beneficiaries to the largest. Figure 1 displays the distribution of beneficiaries attributed to the 54 group practices.

Figure 1. Number of Beneficiaries Attributed to the 54 GPRO Group Practices



In 2011, each beneficiary attributed to a group practice had an average of 10 attribution-eligible E&M visits (both to eligible professionals affiliated with and not affiliated with the group practice), ranging from 6 to 14 visits per group practice. An average of 7 of these E&M visits were with eligible professionals affiliated with the group practice, ranging from 4 to 9 E&M visits. Thus, generally GPRO groups provided not only the plurality but the large majority of attribution-eligible E&M visits to beneficiaries attributed to the group. Figure 2 shows the distribution of the percentage of attribution-eligible E&M visits billed by a group practice. On average, the groups accounted for 76 percent of their attributed beneficiaries' attribution-eligible E&M visits.

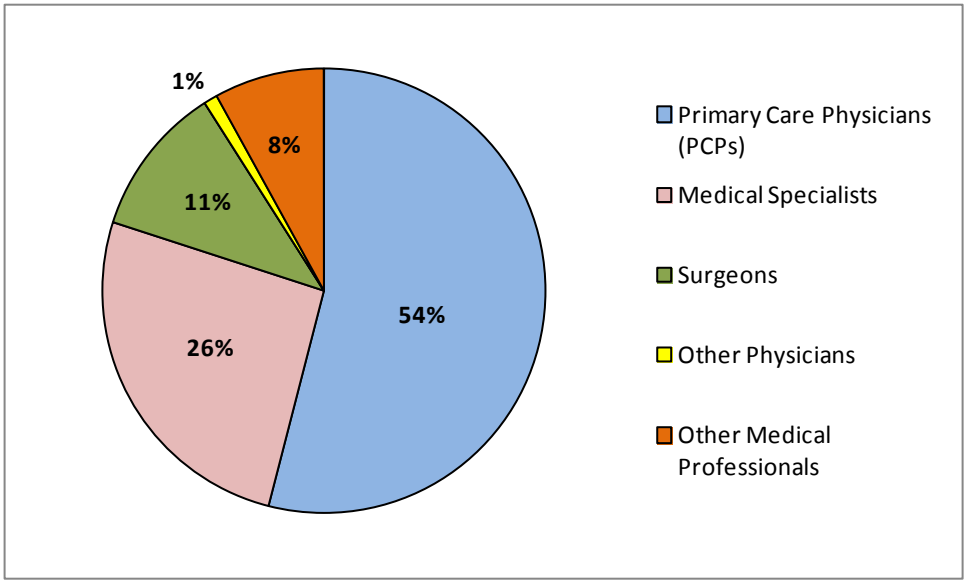
Figure 2. Distribution of the Percentage of All Attribution-Eligible Evaluation and Management Visits Billed by the 54 GPRO Group Practices



In comparing the types of physicians providing the plurality of attribution-eligible E&M visits within a group practice, Figure 3 shows that across the 54 groups, PCPs on average

provided the plurality of E&M visits to more than half (54 percent) of their group's attributed beneficiaries, followed by medical specialists at 26 percent of attributed beneficiaries. Surgeons provided the plurality of E&M visits to 11 percent of beneficiaries on average, other physicians 1 percent, and other medical professionals 8 percent. Across the 54 groups, emergency department physicians on average provided the plurality of E&M services to less than 1 percent of their group's beneficiaries, although for some groups they accounted for the plurality of 1 to 3 percent of beneficiaries. Note that in 11 group practices, medical specialists provided the plurality of E&M care for a higher percentage of their group's beneficiaries than PCPs did.

Figure 3. Physician Broad Stratification Category Billing the Plurality of Attribution-Eligible Evaluation and Management Visits for Beneficiaries Attributed to GPRO Groups, 2011



B. Quality of Care Measures: Clinical Care Measures

As described previously, CMS attributed Medicare FFS beneficiaries to a group practice if eligible professionals affiliated with the group billed for at least two office visits or other outpatient E&M services and the group practice had the plurality of E&M charges for that beneficiary. In 2011, CMS provided to each group a database containing a sample of these attributed Medicare FFS beneficiaries who, based on claims data, met the clinical criteria for 26 specific clinical measures of quality. The group practices were responsible for populating the database to report whether attributed patients had received recommended clinical interventions represented by each of the 26 quality measures. Table 2 displays the quality measures, which consist of National Quality Forum-endorsed quality measures in four disease modules (diabetes mellitus, heart failure, coronary artery disease [CAD], and hypertension), plus four preventive care measures.

Each group practice was required to report clinical data for the first 411 beneficiaries on its list of assigned beneficiaries for each disease module and preventive care measure. If the group practice had fewer than 411 attributed beneficiaries who qualified for a quality measure, clinical indicators had to be submitted for 100 percent of qualified attributed beneficiaries.

Table 4 shows the mean performance rate and the performance rates for the 10th, 50th, and 90th percentiles for each of the 26 quality measures among the group practices. Table 4 also

shows the mean performance rate for 22 comparable measures (based on augmented administrative claims-based measures) that eligible professionals reported at an individual level for the 2011 PQRS program. The mean group practice performance rate was equal to or better than the individual performance rate for 13 of the 22 measures (60 percent), but lower for the other 9 measures. For 7 of the measures, the group mean was at least 10 percentage points higher than the mean individually reported rate.

Table 2. Performance Rates on 26 Quality Measures for GPRO Group Practices, 2011 (percentages)

	Measure Title	2011 Average Performance Rate for Eligible Professionals	2011 Performance Rate for 54 GPROs	Percentile		
			Mean	10th	50th	90th
DIABETES						
GPRO DM-1	Diabetes Mellitus: Hemoglobin A1C Testing	NA	87	74	91	96
GPRO DM-2	Diabetes Mellitus: Hemoglobin A1C Poor Control in Diabetes Mellitus	25	26	15	21	36
GPRO DM-3	Diabetes Mellitus: High Blood Pressure Control in Diabetes Mellitus	76	68	56	69	78
GPRO DM-5	Diabetes Mellitus: Low Density Lipoprotein (LDL-C) Control in Diabetes Mellitus	57	52	38	55	64
GPRO DM-6	Diabetes Mellitus: Urine Screening for Microalbumin or Medical Attention for Nephropathy in Diabetic Patients	80	87	76	89	97
GPRO DM-7	Diabetes Mellitus: Dilated Eye Exam	36	55	34	62	68
GPRO DM-8	Diabetes Mellitus: Foot Exam	56	57	12	60	88
GPRO DM-9	Diabetes Mellitus: Lipid Profile	NA	79	60	83	91
HEART FAILURE						
GPRO HF-1	Heart Failure: Left Ventricular (LVF) Assessment	83	81	60	89	97
GPRO HF-2	Heart Failure: Left Ventricular (LVF) Testing	98	88	73	94	100
GPRO HF-3	Heart Failure: Weight Measurement	NA	75	39	84	93
GPRO HF-5	Heart Failure: Patient Education	50	68	14	79	94
GPRO HF-6	Heart Failure: Beta Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)	96	92	85	95	98
GPRO HF-7	Heart Failure: Angiotenson-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD)	93	86	70	90	98
GPRO HF-8	Heart Failure: Warfarin Therapy for Patients with Atrial Fibrillation	54	78	59	82	93

Table 2 *(continued)*

	Measure Title	2011 Average Performance Rate for Eligible Professionals	2011 Performance Rate for 54 GPROs	Percentile		
			Mean	10th	50th	90th
CORONARY ARTERY DISEASE						
GPRO CAD-1	Coronary Artery Disease (CAD): Oral Antiplatelet Therapy Prescribed for patients with CAD	81	84	56	90	96
GPRO CAD-2	Coronary Artery Disease (CAD): Drug Therapy for Lowering LDL Cholesterol	85	89	76	92	97
GPRO CAD-3	Coronary Artery Disease (CAD): Beta Blocker Therapy for CAD Patients with Prior Myocardial Infarction	84	86	75	89	97
GPRO CAD-7	Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Patients with CAD and Diabetes and/or Left Ventricular Systolic Dysfunction (LVSD)	84	82	71	85	92
HYPERTENSION						
GPRO HTN-1	Hypertension (HTN): Blood Pressure Measurement	0	87	59	97	100
GPRO HTN-2	Hypertension (HTN): Blood Pressure Control	NA	68	57	68	79
GPRO HTN-3	Hypertension (HTN): Plan of Care	82	64	30	71	93
PREVENTIVE CARE AND SCREENING						
GPRO PREV-5	Preventive Care and Screening: Screening Mammography	50	65	47	66	80
GPRO PREV-6	Preventive Care and Screening: Colorectal Cancer Screening	44	58	37	61	74
GPRO PREV-7	Preventive Care and Screening: Influenza Immunization for Patients $\geq$ 50 Years Old	43	64	38	65	82
GPRO PREV-8	Preventive Care and Screening: Pneumonia Vaccination for Patients	55	57	34	58	81

Notes: DM-2 is a measure of poorly controlled blood sugar; higher scores (and percentile rankings) on this measure reflect worse performance.

Within the table, NA indicates that individual eligible professionals did not have an opportunity to report on a PQRS measure that was comparable to the PQRS measure shown for the 54 GPRO group practices.

The GPRO group practice performance rates, which are based on only a sample of the group's attributed beneficiaries, were statistically reliable at a high level across the vast majority

of the measures. In this context, high reliability means a group's performance rates would be highly similar or the same if a different beneficiary sample population of the group practice was used for quality measurement. Although there is no universally agreed upon minimum reliability threshold, reliability scores in the 0.50 to 0.70 range are often considered moderate and scores greater than 0.70 are considered high. The average reliability score for the group practices' clinical quality measures related to CAD ranged from 0.83 to 0.99, for diabetes from 0.95 to 0.99, for heart failure from 0.83 to 1.00, for hypertension from 0.91 to 1.00, and for the preventive measures from 0.97 to 0.99. For all measures, no less than 90 percent of groups achieved a reliability of at least 0.50, with most group practices well above that level.

For quality measures related to CAD, the percentage of groups with performance rates statistically different from the mean (*p*-value less than 0.05) across the 54 groups ranged from 49 to 87 percent. The percentage of groups with performance rates statistically different from the group mean ranged from 57 to 83 percent for diabetes, 48 to 94 percent for heart failure, 61 to 94 percent for hypertension, and 72 to 82 percent for the preventive measures.

We also examined differences in PQRS performance rates between 2010 and 2011 for the group practices that participated in GPRO in both years, and a consistent pattern of differences did not emerge (Table 2a). Of the 26 measures, there were 10 measures with higher average performance rates in 2011. The other 16 measures had lower average performance rates in 2011. The differences were statistically significant for 5 of the 26 measures.

Table 2a. Differences in 2010 and 2011 Performance Rates on 26 Quality Measures for GPRO Group Practices for Groups Reporting Measures in Both Years

	Measure Title	Number of Groups Reporting[a]	2010 Average Performance Rate	2011 Average Performance Rate	Difference (2011-2010)
DIABETES					
GPRO DM-1	Diabetes Mellitus: Hemoglobin A1C Testing	35	93.4	89.6	-3.9*
GPRO DM-2	Diabetes Mellitus: Hemoglobin A1C Poor Control in Diabetes Mellitus	35	22.3	23.9	1.6
GPRO DM-3	Diabetes Mellitus: High Blood Pressure Control in Diabetes Mellitus	35	57.7	69.1	11.4*
GPRO DM-5	Diabetes Mellitus: Low Density Lipoprotein (LDL-C) Control in Diabetes Mellitus	35	54.3	52.6	-1.7
GPRO DM-6	Diabetes Mellitus: Urine Screening for Microalbumin or Medical Attention for Nephropathy in Diabetic Patients	35	89.2	88.1	-1.1
GPRO DM-7	Diabetes Mellitus: Dilated Eye Exam	35	60.6	57.2	-3.4
GPRO DM-8	Diabetes Mellitus: Foot Exam	35	60.8	61.5	0.6
GPRO DM-9	Diabetes Mellitus: Lipid Profile	35	84.3	80.9	-3.4*
HEART FAILURE					
GPRO HF-1	Heart Failure: Left Ventricular (LVF) Assessment	34	86.6	83.6	-3.0
GPRO HF-2	Heart Failure: Left Ventricular (LVF) Testing	34	85.9	88.5	2.6
GPRO HF-3	Heart Failure: Weight Measurement	34	86.3	81.9	-4.4
GPRO HF-5	Heart Failure: Patient Education	34	77.0	73.2	-3.9

Table 2a *(continued)*

	Measure Title	Number of Groups Reporting[a]	2010 Average Performance Rate	2011 Average Performance Rate	Difference (2011-2010)
GPRO HF-6	Heart Failure: Beta Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)	34	92.2	95.1	2.8
GPRO HF-7	Heart Failure: Angiotenson-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD)	34	89.7	89.2	-0.5
GPRO HF-8	Heart Failure: Warfarin Therapy for Patients with Atrial Fibrillation	34	79.8	79.5	-0.3
CORONARY ARTERY DISEASE					
GPRO CAD-1	Coronary Artery Disease (CAD): Oral Antiplatelet Therapy Prescribed for patients with CAD	34	85.4	87.1	1.7
GPRO CAD-2	Coronary Artery Disease (CAD): Drug Therapy for Lowering LDL Cholesterol	34	89.8	91.7	1.9
GPRO CAD-3	Coronary Artery Disease (CAD): Beta Blocker Therapy for CAD Patients with Prior Myocardial Infarction	34	86.8	89.9	3.1*
GPRO CAD-7	Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Patients with CAD and Diabetes and/or Left Ventricular Systolic Dysfunction (LVSD)	34	82.6	83.5	0.9
HYPERTENSION					
GPRO HTN-1	Hypertension (HTN): Blood Pressure Measurement	34	93.2	90.9	-2.3
GPRO HTN-2	Hypertension (HTN): Blood Pressure Control	34	68.4	68.2	-0.2
GPRO HTN-3	Hypertension (HTN): Plan of Care	34	56.1	62.8	6.7
PREVENTIVE CARE AND SCREENING					
GPRO PREV-5	Preventive Care and Screening: Screening Mammography	35	74.5	65.7	-8.8*
GPRO PREV-6	Preventive Care and Screening: Colorectal Cancer Screening	35	59.9	59.8	-0.1
GPRO PREV-7	Preventive Care and Screening: Influenza Immunization for Patients $\geq$ 50 Years Old	35	66.6	65.6	-1.0
GPRO PREV-8	Preventive Care and Screening: Pneumonia Vaccination for Patients	35	62.3	62.0	-0.3

Notes: DM-2 is a measure of poorly controlled blood sugar; higher scores on this measure reflect worse performance.

[a] The number of groups reporting can vary by measure because not every group reported all measures in both years.

* Statistically significant at the $p < 0.05$ level, two-tailed paired t-test.

The percentage of primary care physicians in a group practice did not correlate with the average of the standardized quality measures (correlation of 0.09 and *p*-value of 0.51). Each quality measure was standardized by subtracting the group's performance by the mean performance rate across the 54 groups and then dividing that number by the standard deviation of the performance rate across the 54 groups.

C. Quality of Care Measures: Potentially Avoidable Hospitalization Measures

In addition to the 26 clinical quality measures included in the GPRO PQRS program, the feedback reports contained each group practice's performance on measures of potentially avoidable hospitalizations for ambulatory care-sensitive conditions (ACSCs). These are conditions for which timely outpatient care may prevent complications or more severe disease. The Medicare claims-based measures were derived from Prevention Quality Indicator measures developed by the Agency for Healthcare Research and Quality (AHRQ); more information can be found at http://www.qualityindicators.ahrq.gov/modules/pqi_overview.aspx.

CMS reported ACSCs for diabetes, chronic obstructive pulmonary disease (COPD), heart failure, an acute conditions composite measure, and a total composite measure made up of these four ACSC measures. The diabetes measure is itself a composite measure based on short-term diabetes complications, uncontrolled diabetes, long-term diabetes complications, and lower extremity amputation for diabetes. The acute conditions composite is a combined measure based on bacterial pneumonia, urinary tract infection, and dehydration. The performance rate for each acute condition is computed as the number of hospitalizations for beneficiaries attributed to the group who were identified as having been hospitalized for that condition in 2011 (the numerator), divided by the sum of all beneficiaries attributed to the group practice (the denominator). For the three chronic conditions (diabetes, COPD, and heart failure), CMS calculated the performance rate as the number of hospitalizations for that condition in 2011 (the numerator), divided by the sum of attributed beneficiaries diagnosed with the condition (the denominator). The total composite rate is the sum of the numerators for diabetes, COPD, heart failure, and the acute conditions composite divided by the sum of the corresponding measure denominators. For all ACSC measures, the performance rates are expressed per 1,000 attributed beneficiaries. The ACSC measures were not risk-adjusted for beneficiary demographic characteristics or disease status. Table 3 shows the 54 groups' mean, minimum, and maximum performance rates for each of the ACSC measures. The range in group performance was substantial across all ACSC measures.

Table 3. 2011 Performance Rates (hospitalizations per 1,000 attributed beneficiaries) for the Ambulatory Care Sensitive Conditions for the 54 GPRO Group Practices

Measure	Mean	Minimum	Maximum
Total Composite	38	20	82
Diabetes	27	9	57
Chronic Obstructive Pulmonary Disease	140	71	332
Heart Failure	111	51	201
Acute Conditions Composite	24	12	43

For the ACSCs related to chronic conditions, most group practices achieved a reliability score greater than 0.70. In particular, 52 of 54 groups for diabetes, 53 of 54 groups for COPD, and 53 of 54 groups for heart failure achieved a reliability score greater than 0.70. Of those group practices that did not have reliability scores greater than 0.70, all achieved statistical reliability of greater than 0.50. For the acute conditions composite, 52 of 54 groups achieved a reliability score greater than 0.70 and all groups achieved statistical reliability greater than 0.50. All groups achieved a reliability score greater than 0.70 for the total composite ACSC rate.

For each of the ACSCs related to chronic conditions, more than half of group practices reported rates statistically different from the mean (*p*-value less than 0.05) across the 54 groups. For the acute conditions composite, 57 percent of groups reported a rate statistically different from the group mean.

D. Quality of Care Measures: Hospital Discharge Measures

In addition to ACSC measures, CMS reported GPRO group performance on 30-day hospital post-discharge provider visits and all-cause hospital readmissions. Table 4 shows the 54 groups' mean, minimum, and maximum performance rates for each measure. Across the 54 groups, average performance on 30-day post-discharge provider visits was 788 per 1,000 discharges. The average all-cause 30-day readmission rate was 154 per 1,000 discharges. The range in group performance was fairly large for the all-cause 30-day readmission rate.

Table 4. Performance Rates (follow-up visits or hospital readmissions per 1,000 discharges) for the Hospital Discharge Measures Among the 54 GPRO Group Practices

Measure	Mean	Minimum	Maximum
Physician Follow-Up Visit Within 30 Days of Discharge	788	702	844
All-Cause 30-Day Readmission Rate	154	103	266

For the 30-day post-discharge provider visit measure, 53 of 54 groups achieved a reliability score greater than 0.70. The remaining group achieved a reliability score slightly less than 0.50. For the all-cause 30-day hospital readmissions measure, 53 of 54 groups achieved statistical reliability greater than 0.70 and all groups achieved statistical reliability greater than 0.50.

For the 30-day post-discharge provider visit measure, 72 percent of groups reported rates statistically different from the mean, whereas 65 percent of groups reported rates statistically different from the group mean for all-cause 30-day hospital readmissions.

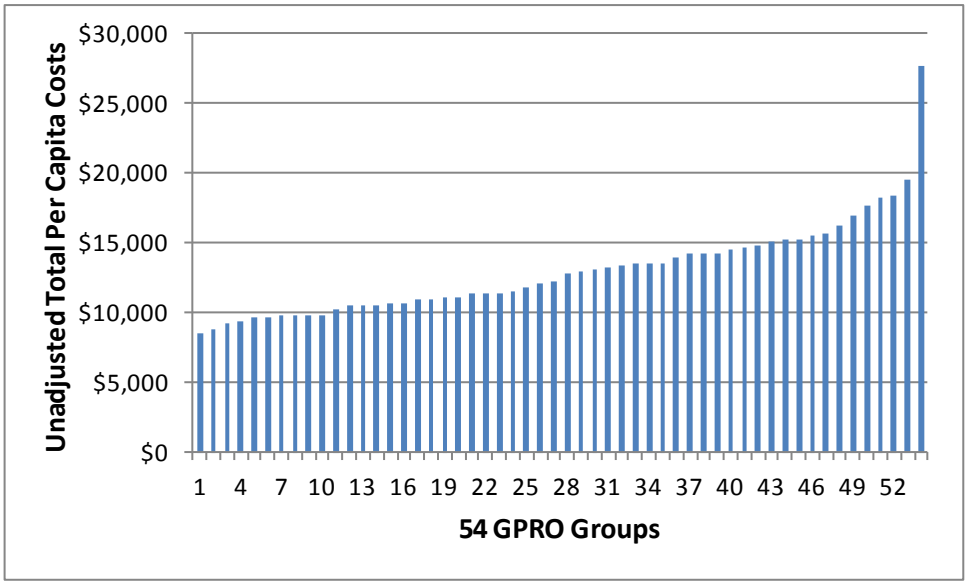
E. Resource Use Measures: Per Capita Cost Measures

Five measures of group resource use were examined: total per capita costs for beneficiaries attributed to the group practice and per capita costs for beneficiaries attributed to the group who had one of the following four chronic conditions: diabetes, heart failure, COPD, and CAD. In calculating these measures, CMS first standardized Medicare payments for geographic and other price differentials to ensure fair comparisons among groups. Geographic variations in Medicare payments to providers can reflect factors unrelated to the care provided to beneficiaries. All

Medicare payments have been standardized such that a given service is assigned the same dollar value across all providers within the same facility type or setting, regardless of geographic location or differences in Medicare payment rates among facilities. More information about how CMS standardized payments can be found in the document describing the methodologies used in the 2011 Quality Resource and Use Reports (QRURs), which can be accessed at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/downloads/2011_group_detail_methodology.pdf.

Across the 54 group practices, the average payment-standardized (but nonrisk-adjusted) total per capita costs for attributed beneficiaries was $12,997. Figure 4 displays the range of total per capita costs from lowest to highest: costs ranged from $8,539 to $27,618, for a total per capita difference of $19,079.

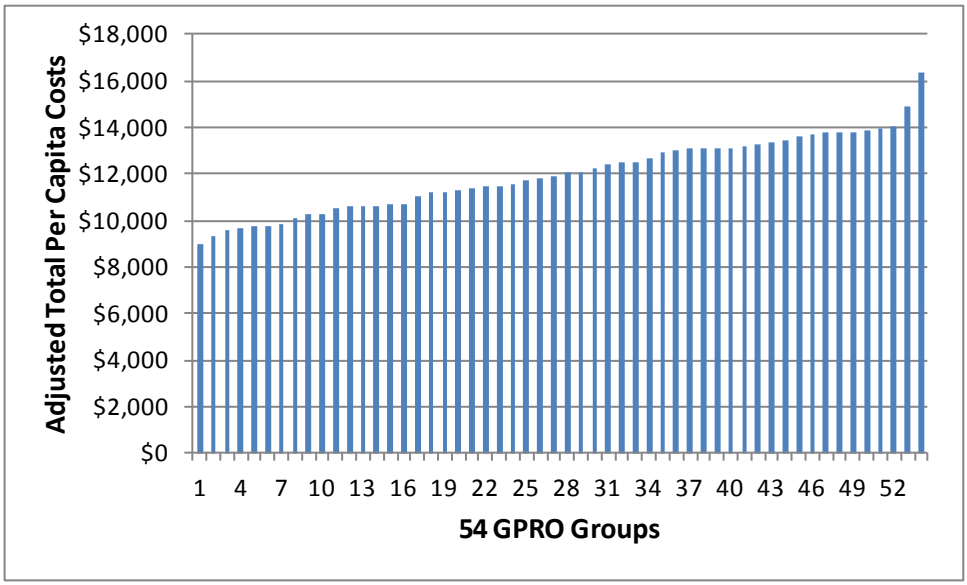
Figure 4. Payment-Standardized but Nonrisk-Adjusted 2011 Total Per Capita Costs for the 54 GPRO Group Practices



After payment-standardizing per capita cost measures, the measures were also risk-adjusted to account for the unique mix of Medicare beneficiaries attributed to each group. Recognizing that physiologic differences among beneficiaries can affect medical costs regardless of the care provided, the risk adjustment methodology includes markers for patient demographics, socioeconomic factors, and prior medical diagnoses. The per capita cost measure risk-adjustment methodology is based on CMS' hierarchical condition categories (HCC) model that assigns International Classification of Diseases, Ninth Edition (ICD-9) diagnosis codes (each with similar disease characteristics and costs) to 70 clinical conditions to capture medical condition/cost risk. The prior-year HCC risk scores used in the per capita cost risk-adjustment model also incorporate patient age, gender, reason for Medicare eligibility (aged or disabled), and Medicaid eligibility; the model also accounts for whether a beneficiary was diagnosed in the previous year with end-stage renal disease (ESRD). More information about CMS risk-adjusted per capita costs can be found in the document describing the methodologies used in the 2011 QRURs, which can be accessed at http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/PhysicianFeedbackProgram/downloads/2011_group_detail_methodology.pdf.

Figure 5 shows that, after risk-adjustment, the average adjusted total per capita costs across the 54 groups were $11,983, with a range of $8,989 to $16,353, for an overall difference of $7,364. Thus the risk-adjustment methodology had the effect of reducing the absolute difference between the groups with the lowest and highest total per capita cost by 61.4 percent. In particular, the lowest-cost one-third of the groups was adjusted upward by an average of 3.5 percent and the highest (most expensive) one-third of the groups was lowered by 16.9 percent. The middle third of the groups, on average, had per capita costs adjusted downward by 2.5 percent, with the range of these adjustments being –16.1 to +6.2 percent.

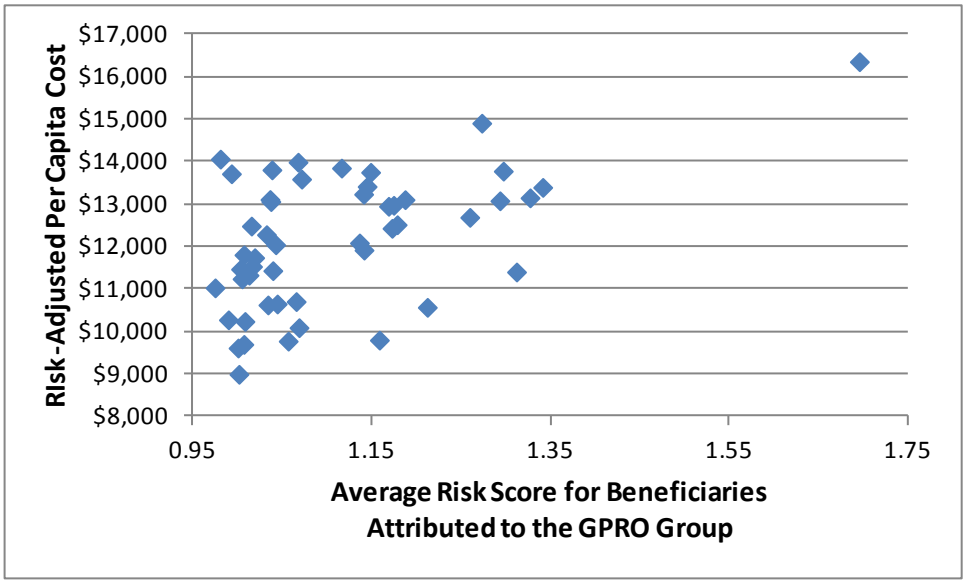
Figure 5. Payment-Standardized, Risk-Adjusted 2011 Total Per Capita Costs for the 54 GPRO Group Practices



Among the 11 groups for which medical specialists provided the plurality of care to attributed beneficiaries, 10 medical groups had their costs risk-adjusted downward. Among groups for which PCPs provided the plurality of care to attributed beneficiaries, the groups were nearly evenly split between those who had their costs risk-adjusted upward or downward.

Figure 6 shows the relationship between a group's average risk and its risk-adjusted total per capita costs. Although there is a positive correlation (0.59), risk-adjusted total per capita costs are still fairly dispersed at any given level of risk.

Figure 6. Relationship Between Risk-Adjusted Total Per Capita Costs and Average HCC Risk Score for the 54 GPRO Group Practices, 2011



The 2011 feedback reports also showed the percentage of eligible professionals who treated the beneficiaries attributed to the group practice but who did not bill under the group practice's TIN.[1] No meaningful association was found (correlation of 0.12) between the percentage of professionals who did not bill under the group practice's TIN and higher total per capita costs for the group's attributed beneficiaries.

All 54 group practices achieved statistical reliability scores greater than 0.70 for the total per capita cost measures, as did 53 of the 54 group practices for the four chronic condition-specific cost measures. The group practices achieved an average reliability score of 0.99 for the total per capita cost measure. Average reliabilities for the condition-specific cost measures were 0.93 for heart failure, 0.89 for COPD, 0.95 for diabetes, and 0.95 for CAD.

Nearly 90 percent (48 of 54) of group practices reported total per capita costs that were statistically different from the group mean at the 5 percent level of significance. However, the percentage of group practices with condition-specific per capita costs statistically different from the mean was lower across all conditions, at 59 percent for heart failure, 56 percent for COPD, 70 percent for diabetes, and 70 percent for CAD.

1. Comparison of Quality of Care to Cost of Care

A simple quality composite score was constructed by combining the 26 clinical quality measures, the chronic conditions ACSC composite[2] and acute conditions ACSC composite, and

[1] On average, 45 percent of the eligible professionals that billed at least one Part B professional claim line item for attributed patients were outside the group practice. The portion of eligible professionals seeing attributed beneficiaries but not billing under the group TIN ranged from 16 to 83 percent.

[2] The chronic conditions composite was constructed as the sum of the numerators for diabetes, COPD, and heart failure ACSC measures divided by the sum of their corresponding denominators.

the two hospital discharge measures. The quality composite score was computed by first standardizing each measure by subtracting the group's performance rate by the mean performance rate across the 54 groups and then dividing that number by the standard deviation of the performance rate across the 54 groups. Then the following standardized measures were multiplied by –1 so that higher rates corresponded to better performance: DM-2, chronic conditions ACSC composite, acute conditions ACSC composite, and all-cause hospital readmissions. Each measure was then assigned to one of three domains. All 26 clinical quality measures except for PREV-7 (Influenza Immunization) were assigned to the Clinical Process/Effectiveness domain; PREV-7 was assigned to the Population/Public Health domain; and the ACSC composites and hospital discharge measures were assigned to the Care Coordination domain. For each group, the simple average was computed across its standardized scores included in each domain. Finally, for each group, the simple average across the three domain scores was computed to arrive at the group's quality composite score.

Figure 7 is a scatter diagram that displays the relationship between the composite quality score for each group practice and the total payment-standardized risk-adjusted per capita cost measure. Although there is a negative correlation (-0.53), total per capita costs are fairly dispersed at any given level of quality.

Figure 7. Relationship Between Quality Composite Score and Risk-Adjusted Per Capita Costs for the 54 GPRO Group Practices, 2011