

The Centers for Medicare & Medicaid Services' Office of Research, Development, and Information (ORDI) strives to make information available to all. Nevertheless, portions of our files including charts, tables, and graphics may be difficult to read using assistive technology.

Persons with disabilities experiencing problems accessing portions of any file should contact ORDI through e-mail at ORDI_508_Compliance@cms.hhs.gov.

University of Minnesota

*Twin Cities Campus
Division of Health Policy and Management
School of Public Health
420 Delaware Street S.E.
Minneapolis, MN 55455*

Alternative Approaches to Measuring Physician Resource Use

**Medicare/Medicaid Research and Demonstration
Task Order Contract (MRAD/TOC)
HHSM-500-2005-000271 Task Order 0004**

Interim Report

March 2010

Submitted by

**David Knutson, M.S.
Bryan Dowd, Ph.D.
Robert Kane, M.D.**

CMS Project Officer: Craig Caplan

Abstract

The goal of this project is to explore potential measures that would support public reporting and a payment system that would reward physicians in traditional fee-for-service (FFS) Medicare for efficient and high “quality” care. The project is premised on the idea that physicians can be held accountable for the costs and quality of care received by their patients, but such accountability requires a reliable system for attributing care to providers. Our approach to attribution recognizes that there are two types of physicians – those who primarily provide care for hospitalized beneficiaries (referred to as Medicare Severity Diagnosis Related Group or MS-DRG approach) and those who provide primarily outpatient care (referred to as the beneficiary-level approach). We develop different approaches to attribution and different quality measures for these types of physicians. The MS-DRG approach tests several possible accountability algorithms based on the attending, operating, and performing physician. For a selected set of conditions, we found that 30 and 60 day windows captured most of the costs associated with inpatient and follow-up care (excluding the MS-DRG payment itself, since it is prospective and fixed, given that the patient is admitted). Including severity measures in the risk adjustments did not add significantly to the explained variance in our cost equations. Accountability in the beneficiary-level analysis is based on an analysis of CMS’s Physician Quality Reporting Initiative (PQRI) reporting system. The beneficiary-level analyses suggested that even though initial PQRI reporting rates are low, PQRI-reported beneficiaries are remarkably similar to non-reported beneficiaries seeing the same physicians.

Introduction

The goal of this project is to explore potential measures that would support public reporting and a payment system that would reward physicians in traditional fee-for-service (FFS) Medicare for efficient and high “quality” care. The activities addressed in this project are premised on the idea that physicians can be held accountable for the costs of care received by their patients and the quality associated with that care. There is little point in developing elaborate performance measures of efficiency and quality if the care received (or not received) by a beneficiary cannot be attributed to a particular health care provider or set of providers. Thus, the first problem that must be addressed is the attribution of beneficiaries to providers.

In a health insurance program like FFS Medicare, such attribution is difficult. Patients are not restricted in their choice of physicians. There is no enforced system of primary care. The structure of the program invites uncoordinated care.

Our approach to attribution recognizes that there are two types of physicians – those who primarily provide care for hospitalized beneficiaries and those who provide primarily outpatient care. Our measures of subsequent use of resources for physicians providing inpatient care are based on starting with Medicare Severity Diagnosis Related Groups (MS-DRGs). The MS-DRG analysis is conducted *conditional* on the patient having been admitted. As discussed in this report, attribution of inpatient medical care to the attending physician is relatively straightforward, but attribution of inpatient surgical care to the operating versus performing physician requires analysis that we are currently performing.

Our approach to physicians providing primarily outpatient care is to explore the usefulness of data from CMS’s Physician Quality Reporting Initiative (PQRI) for attribution purposes. Our analysis of physicians providing primarily outpatient care is referred to as the beneficiary level analysis. The beneficiary level analysis incorporates measures of resource use and quality that involve the use of both inpatient and outpatient services (e.g., PQRI-based measures of appropriate outpatient care and measures of appropriate use of emergency departments and the hospital).

In addition to our work on the attribution problem, we have begun to develop measures of physician resource use in the MS-DRG analysis. All measures of resource use must be risk-adjusted, and the measures of resource use and quality must be combined to produce metrics that can be used to reward efficient, high-quality physician practices. This report includes the results from analyses of risk adjustment in the MS-DRG analysis, and briefly describes our approaches to combining measures of resource use with measures of quality.

Both the MS-DRG and beneficiary level analyses are in the initial stages, and we are not in a position to make final conclusions regarding those analyses. This report focuses to a large extent on descriptive data and anomalies, inconsistencies, and other features of the data encountered thus far in the analyses.

The remainder of the report is organized as follows:

- I. Background and Policy Context
- II. Project Overview
- III. Interim Findings from Measure Development and Testing: MS-DRG-Based Approach
- IV. Interim Findings from Measure Development and Testing: Beneficiary level Approach
- V. Next Steps
- VI. Conclusion

I. Background and Policy Context

Health care purchasers have long been aware of variation in the amount and the quality of care delivered in the United States. The relationship between quality of care and the cost of care remains variable; one cannot assume that more care is better care or that more money automatically buys better care.

Physician care is an important source of the cost growth in the FFS Medicare program. Attempts by Congress to establish across-the-board spending limits through the Sustainable Growth Rate legislation have been largely unsuccessful.

In the *Medicare Improvements for Patients and Providers Act (MIPPA) of 2008* Congress has directed the Centers for Medicare and Medicaid Services (CMS) to: “develop a plan to transition to a value-based purchasing program for Medicare payment for physician and other professional services.” In response, CMS is seeking to understand and develop mechanisms to assess the variation in resource use among practitioners, and the relationship between resource use and quality. This information could serve as a basis for assessing the value of care by providers.

The statement of work for this project was as follows (footnotes deleted):

“CMS and others in the policy community have been increasingly focused on moving to a value based purchasing (VBP) system for traditional Medicare, in which physicians’ payments would differ depending on the “value” of services they provided. Under VBP, physicians who routinely use relatively few resources while maintaining adequate quality would receive larger payment updates than other physicians. The Medicare Payment Advisory Commission (MedPAC), the U.S. Government Accountability Office (GAO), and others have issued reports on this subject. While much of the work, including that by CMS, has focused on the use of commercially developed episode groupers to measure resource use, little peer reviewed literature that evaluates these groupers exists, and few studies have examined feasible alternatives in Medicare. The purpose of this solicitation is to suggest and develop alternative approaches to the commercial episode grouper-based resource use measures.”

When we initiated this project there was already considerable relevant work underway to adapt techniques for measuring physician resource use that had been developed in the commercial sector to the context of Medicare. However, Medicare beneficiaries differ from those in the commercial sector by having more medical conditions (and hence using more physicians). Whereas the modal commercial enrollee may only one chronic illness, the typical Medicare beneficiary likely will have several chronic conditions and see many physicians. As a result, attributing utilization and quality is much more difficult, and the techniques in the commercial sector consequently are of limited value for FFS Medicare.

CMS first undertook a review of two common commercially used tools, Episode Treatment Groups (ETGs) and Medical Episode Grouper (MEG), and assessed their potential for applicability to care delivered to Medicare beneficiaries.¹ While showing some applicability, these tools encounter challenges in the Medicare setting. These challenges include:

- The payment methods across care settings under Medicare.
- The increased prevalence of multiple chronic conditions in Medicare beneficiaries.
- The greater number of practitioners caring for each Medicare beneficiary.

CMS continues to explore and consider approaches to physician resource use measurement using the tools described above. This project explores additional alternatives using CMS administrative data and information from previous CMS-sponsored research.

II. Project Overview

This project is sponsored by CMS under a Task Order Contract and is scheduled to conclude in mid-2011. This Interim Report presents progress to date. This project investigates not only new methods of evaluating resource use, but also how measures of resource use can be coupled with measures of quality, providing a more comprehensive perspective of value.

Resource use can be measured on a per episode basis or a *per capita* (or beneficiary) basis. *Episode-based* resource use measurement examines care services over a defined time period indexed to specific conditions or care events. Our MS-DRG analysis, designed to measure resource use and quality for physicians primarily providing inpatient care, is an example of an episode-based approach. Our *per capita* or beneficiary level analyses focus on all expenditures for a beneficiary incurred during a defined time period.

Our **episode of care approach** is based on MS-DRGs and subsequent post-acute care, including readmissions. This approach focuses on the accountability of those physicians who work in hospitals and is compatible with other payment reform initiatives seeking to bundle

¹ MaCurdy, T., Kerwin, J., Gibbs, J., et al.: Evaluating the Functionality of the Symmetry ETG and Medstat MEG Software in Forming Episodes of Care Using Medicare Data. Acumen, LLC. Burlingame, CA. August 2008a. Internet address: <http://www.cms.hhs.gov/Reports/downloads/MaCurdy.pdf>. (Accessed December 1, 2009.)

payments and accountability for hospital and post-hospital care.

Our **beneficiary per capita approach** holds a physician or physician group accountable for an entire year of care for a beneficiary. The purpose of the beneficiary level analysis is to develop cost and quality data on primary care physicians who treat beneficiaries primarily in ambulatory care settings. However, since all care during the year is included, the physician will be held responsible for the cost of hospitalizations as well as the appropriateness of hospitalizations and emergency department utilization as part of the quality measures.

A. Attributing the care of beneficiaries to physicians

In order for the development of any measure of resource use or quality to be useful, that development must be preceded by a system for attributing the care of beneficiaries to physicians who are accountable for that care. Attribution is an issue that applies equally to the MS-DRG analysis for physicians providing inpatient care and the beneficiary level analysis for physician providing outpatient care.

There are two dimensions of attribution. The first distinguishes between “active” attribution that is initiated by an agreement between the physician and patient, versus “passive” attribution initiated purely by the analyst/payer, often without the physician’s prior knowledge or consent. The second dimension distinguishes between *ex ante* accountability in which the physician is linked to a patient at the beginning of the observation period (e.g., a year) versus a *post hoc* approach that links the physicians to patients based on events that occur during the year.

1. Attribution in the MS-DRG analysis

We have tested alternative ways to find the appropriate hospital-based physician to whom the MS-DRG-based episodes could be attributed. In each case, we have identified physicians by their National Provider Identifiers (NPIs).

The attribution methods were dependent on whether the MS-DRG was a medical or a surgical case. For medical cases, the answer was clear-cut: we used the attending physician listed on the Medicare Part A claim as the clinician primarily responsible for the care of the patient from the beginning of the hospital episode. For surgical cases, we compared the consistency of three identification approaches:

1. The attending physician: the clinician primarily responsible for the care of the patient from the beginning of a hospitalization (medical or surgical case).
2. The operating physician listed on the Medicare Part A claim: the clinician performing the principal procedure for surgical MS-DRGs.
3. The physician performing the principal procedure according to the associated Part B claims.

The match between Part A and Part B physicians was lower than expected; however, the

congruence between the operating physician and the performing physician was better than the congruence between the attending and performing physician for surgical DRGs. Hence, we have decided to create and test two separate physician profiles for surgical MS-DRGs, one where we attributed costs to the Part A operating physician and the other where we attributed costs to the Part B performing physician. Those tests are currently in progress.

2. Attribution in the beneficiary level analysis

Using 2008 data from Colorado, we are testing CMS's PQRI system as a method of attributing patients to physicians. The PQRI program is CMS's initial venture into paying physicians for reporting in the Medicare FFS system. PQRI began in 2007 and is a voluntary program in which physician practices can receive payment by reporting on a subset of defined PQRI quality measures. The quality measures are most frequently reported on special fields of the CMS 1500 billing form or through certified registries. In 2009, about 25-30 percent of US physician practices (defined by rolling up reporting NPIs to the tax identification number (TIN) level) reported PQRI quality information.

In the PQRI program, physicians choose to report a PQRI quality measure for a patient, an action that we interpret as being a proxy for self-attribution of patients by the physicians. While this is not a comprehensive attribution method, as is the case for medical homes or for an accountable care organization (ACO) model, PQRI reporting may be viewed as implying an active taking of responsibility that cannot be inferred from passive attribution methods. Thus, our beneficiary level approach may provide the opportunity to both improve attribution acceptance in the current FFS system and develop a novel beneficiary level efficiency measure that, while designed to function in the current system, also could be used in an ACO-based system. As reported in Section IV, we find that despite low levels of reporting in the nascent PQRI reporting system, measures of patient "adherence" or "loyalty" to the PQRI reporting physician compare favorably with the upper bound of such measures based on passive attribution systems (e.g., the physician with the maximum number of visits for each beneficiary).

B. Data Sources

For the initial stage of measure prototype development, we used claims and enrollment data for FFS beneficiaries in Colorado. Colorado simply represents a manageable convenience sample of beneficiaries and providers with a mix of socio-demographics and types of physician practices. The data sample for our initial MS-DRG analyses included beneficiaries in Colorado with a 2008 inpatient claim with any of the MS-DRG codes associated with the selected MS-DRGs. Our beneficiary level PQRI analysis used a 100 percent sample of 2008 data reported by physicians in Colorado. We obtained PQRI quality reporting information directly from the claims which were submitted by physicians.

C. Measure Development Steps

The following are steps being taken in the development of the measures in this project. The steps are progressing in parallel for the episode and the beneficiary level approaches. There will be overlapping elements in later steps, such as combining cost performance measures with

quality measures and overuse measures.

1. Create Medicare **episodes of care** based on selected MS-DRGs.² These acute illness episodes include inpatient stay(s), outpatient procedures, and scheduled outpatient services as well as associated post-acute care costs and other costs that are clearly associated with the episode.
2. Create a **beneficiary level cost of care** approach that builds on PQRI reporting wherein costs are potentially assigned to the beneficiary-year without attempting to assign costs to specific conditions.
3. Compute the risk-adjusted average cost of care for physician practices for each approach. The Hierarchical Condition Category (HCC) model that is used for Medicare Advantage payment under Medicare is adapted for each MS-DRG (to recognize that hospitalized patients are likely sicker than the average beneficiary); in addition, we apply a severity measure based on prior utilization (i.e., did the patient have a previous admission for the same condition in the year prior to this admission). For the beneficiary level approach, we use the HCC based risk adjuster and adjust for beneficiaries' costs that are clearly not attributable to ambulatory care physicians (e.g. initial trauma care).
4. Identify measures of clinical evidence-based performance for each of the two types of physician practice settings – hospital and ambulatory care.
5. Combine the cost and clinical performance measures into measures that can be used to reward efficient, high-quality physician practices.
6. Evaluate the fairness, meaningfulness, and potential to guide improvement of the measures, once developed, with clinical technical expert panels.
7. Reproduce measures on a representative regional sample of US physician practices to evaluate regional variation.

We have initiated analyses one through three and the results are part of this interim report. The results of analyses four through seven will be included in future reports.

² The conditions include acute myocardial infarction, congestive heart failure, stroke, hip fracture, hip replacement, knee replacement, cholecystectomy, pneumonia/bronchitis, chronic obstructive pulmonary disease, back pain, breast cancer, lung cancer, colon cancer, and acute myelogenous leukemia.

III. Interim Findings from Measure Development and Testing: MS-DRG-Based Approach

A. Cost of Care Measure Development and Testing

We began the MS-DRG analysis of physicians providing primarily inpatient care by creating MS-DRG-based cost of care measures.³ This approach attributes to the hospital physician of record the costs of all care within a time window after discharge and associated charges during the hospitalization.

We did not attribute the MS-DRG Part A payment for the initial hospitalization to the physician because the admission was not necessarily in his/her control. Furthermore, that payment is paid prospectively and *once the beneficiary is admitted to the hospital*, that component of cost is not influenced by physician behavior. However, we assumed that hospital readmissions within the chosen time window were attributable to the appropriate hospital physician, as discussed in Section II.

For the prototype measure development, we identified a set of MS-DRGs on the basis of their high frequency and cost for the Medicare program. Our goal was to create a prototype measurement approach that encompassed risk-adjusted cost of care for 30- and 60- day windows of the selected list of conditions by their MS-DRGs in a way that is acceptable to physicians and easily interpreted. The conditions selected were: acute myocardial infarction, congestive heart failure, stroke, hip fracture, hip replacement, knee replacement, cholecystectomy, pneumonia/bronchitis, chronic obstructive pulmonary disease, back pain, breast cancer, lung cancer, colon cancer, and acute myelogenous leukemia. Other conditions could of course have been chosen, but we knew of no reason that this particular mix of conditions would be misleading or unrepresentative. To test this, we have refined the approach with a selected subset of additional MS-DRGs (Medical AMI, AMI + PTCA, Hip Replacement, Knee Replacement, and COPD) and will extend our work to the remaining conditions. As noted in the previous section, the data sample for our initial work included beneficiaries in the state of Colorado with a 2008 inpatient claim with any of the MS-DRG codes associated with the selected MS-DRGs.

In our initial analysis, we excluded beneficiaries who had admissions in the first and last quarters of 2008. For each of the remaining beneficiaries, we calculated the total Medicare costs incurred from the date of admission (i.e., the first day of an index inpatient hospital admission for a related MS-DRG) to the end of the year. We then examined costs incurred from the date of admission to 30, 60, and 90 days after the admission and costs from 30, 60, and 90 days after admission to the end of the year.

All Medicare Part A and Part B costs were included in this analysis, except for MS-DRG Part A costs. We did retain the operational outlier payment amount associated with the initial

³ Although the MS-DRG and associated post-acute care can be considered an “episode” of care, we try whenever possible to use the term MS-DRG rather than “episode” in order to avoid confusion with other episode grouper approaches that are being developed in other projects.

MS-DRG excluding the capital component. Finally, to make hospital readmission costs independent of hospital type, we excluded the operational disproportionate share and operational indirect medical education amounts from all hospital readmissions.

We compared the percentage of these costs accounted for by considering 30-, 60-, and 90- day windows. The results of these analyses showed that 45 to 65 percent of Medicare costs were incurred during the 30-day windows, 60 to 80 percent were incurred during the 60-day windows, and 70 to 85 percent were incurred during the 90-day windows. Because the 90-day windows did not offer much gain over the 60-day windows in terms of explaining additional Medicare costs, we created physician profiles for only 30- and 60- day windows.

We created new risk adjustment weights by adapting the HCC model that is used for Medicare Advantage payments under Medicare. This model assigns a number of major diagnoses to condition categories. The condition categories serve as binary variables in a regression-based risk model. CMS estimated models to predict total annual expenditures for Medicare Advantage beneficiaries using a sample of all Medicare FFS beneficiaries. Because our predicted cost variable is related only to hospitalizations, these HCC weights would not apply. All diagnoses occurring 12 months prior to a given index inpatient admission for beneficiaries were used to count HCCs. The independent variables used in our risk adjustment model included:

- 70 HCC diagnostic categories
- 24 Age/Sex groups
- Medicaid interacted with sex and age/disabled entitlement status
- Originally disabled status interacted with sex and MS-DRG status

We also added severity adjustments based on prior hospitalizations or ER use for the same condition in the 12 months before the admission, since the occurrence of these prior events signals a more severe case than having the diagnosis without those events. We ran risk adjustment models for adjusting 30- and 60- day costs for beneficiaries with admissions for Total Hip Replacement and COPD because the number of beneficiaries with these two conditions offered adequate sample sizes. Adding the two severity variables produced only a slight improvement (0.5 percent-1.0 percent) in the model fit for COPD but did not affect the model for Total Hip Replacement. We will continue to explore the value of adjusting for severity because it may provide clinicians with a greater sense of fairness about the process.

B. Future directions in the development of the MS-DRG cost of care measure

One problem in our research using Colorado data is that 40 percent of physicians had only a single MS-DRG-based episode of care in their profile for each of the MS-DRGs we used. The percentage of low-volume (less than 10 cases) physicians was greater for medical MS-DRGs compared to surgical ones. A high proportion of low-volume (i.e., less than 10 cases per DRG) physicians would considerably reduce the number of physicians for whom reliable physician profiles could be generated for development of measures of physician performance.

One potential partial solution to the small sample problem is to aggregate across physicians. Medicare payments are made to a physician practice based on TIN – that is, the TIN is the unit of accountability in the CMS payment context. Many NPIs can be assigned to a single TIN. Aggregating NPIs by TINs may reduce the proportion of low-volume physicians, but this approach may not entirely solve the problem. Alternatively, aggregation can take place across multiple non-related MS-DRGs attributable to the same physician or TIN. We currently are testing both aggregation methods.

A second problem was that many of the claims showed questionable dollar amounts. This could happen when the Medicare costs seemed to represent only a portion of the full costs (perhaps because another payment source was involved). Beneficiaries in our sample had continuous Part A and Part B enrollment. We excluded cases where Medicare was not the primary payer of the initial MS-DRG hospitalization. However, we did not exclude cases where there was a secondary Part B payer (e.g., the beneficiary had Medicaid or Medigap). These low-cost episodes also skew the cost data that are part of our performance estimates. In an effort to preserve as many cases as possible, we are using an approximation of standardized costs in all our analyses to fix the problem of questionable dollar amounts on claims.

Finally, we need to continue to refine the risk adjustments in our estimates. We will expand the risk adjustment analyses from Colorado to data from 10 states. This will increase the sample size for most of the MS-DRGs and also make the risk adjustment impartial to regional practice patterns.

IV. Interim Findings from Measure Development and Testing: Beneficiary level Approach

A. Cost of Care Measure Development and Testing

The purpose of the beneficiary-level analysis is to develop cost and quality data on physicians who treat beneficiaries primarily in ambulatory care settings. The first step is to develop a methodology for attributing beneficiaries to physicians and this has been the focus of our work to date. We are testing CMS's PQRI system as an attribution methodology. Although the PQRI system has been in place for several years, we are not aware of any rigorous analysis of the system or its potential usefulness for establishing accountability.

The PQRI reporting rules require that a physician report an NPI on the claims form carrying the quality performance data. PQRI reporting is voluntary. Under the current rules, physicians are rewarded simply for reporting PQRI data. There are no rewards or penalties tied to actual process-of-care or health outcome results. Although we refer to the PQRI-reporters as "physicians," PQRI data actually are reported by NPIs, which may represent anything from a family practitioner to a cardiologist to a grocery store. We currently are using provider specialty codes to understand what types of organizations are reporting the various PQRI measures. Given our emphasis on outpatient primary care, we have limited our analyses to PQRI claims that are attached to evaluation and management (E&M) claims for non-hospital services.

The PQRI analysis is based on a 100 percent sample of 2008 data reported by physicians in Colorado. Our findings to date include the following descriptive statistics:

- A prerequisite for any attribution system based on non-hospital E&M PQRI reporting is widespread participation by physicians. Of the 31,210 NPI codes that billed for a non-hospital E&M visit in Colorado in 2008, only 3.9 percent reported a non-hospital E&M PQRI measure on a beneficiary. Internal medicine and cardiologist physicians report PQRI measures on the greatest number of beneficiaries. Although the PQRI system is in its earliest stages of development, these participation rates would need to improve dramatically in order to base an efficiency and quality measurement system on PQRI attribution.
- The results based on beneficiaries, rather than physicians (NPIs), were somewhat better. Of the 87,151 beneficiaries in the 2008 Colorado data, 31 percent had some type of PQRI report filed by a physician. However, only 11.5 percent had an E&M PQRI report and only 9.6 percent had a non-hospital E&M PQRI (i.e., not associated with a hospital stay).
- If a beneficiary receives a PQRI report from more than one physician, it would be necessary either to develop a rule to allocate the beneficiary to a single physician or to apportion the care between multiple reporting physicians. However, multiple physicians reporting on the same beneficiary is not yet a serious problem in our 2008 Colorado data, due in part to the low reporting levels. Of the NPIs reporting E&M PQRI measures in the ambulatory setting, 88 percent were reported by only one NPI.
- The beneficiaries who had a PQRI report were of similar age, race, and sex to non-reported beneficiaries, but the reported beneficiaries tended to be slightly older.
- We were interested in how physicians choose the beneficiaries on which to file PQRI reports. Are they beneficiaries who use more services, for example, recognizing that a slightly higher rate for PQRI-reported patients should be expected, given that PQRI-reported beneficiaries (by definition) have medical conditions that make them eligible for PQRI reporting? We found that beneficiaries with non-inpatient E&M PQRI reports averaged 12 non-inpatient E&M visits during the year, while non-reported beneficiaries who had at least one ambulatory E&M claim from the same set of physicians averaged 11 visits per year.
- Next, we compared the number of visits for reported and non-reported beneficiaries within broad diagnostic categories. For diabetic beneficiaries, the number of visits for reported versus non-reported beneficiaries seeing the same physician is 13.2 versus 12.9. The numbers of visits for beneficiaries with CHF were 27.5 compared to 28.0 for non-reported beneficiaries.

- Do physicians appear to be “cherry-picking” beneficiaries on whom to report PQRI measures? To test that hypothesis we computed concurrent (2008) HCC risk scores for each beneficiary in our condition-specific analysis using the Physician Group Practice Demonstration project’s concurrent coefficients. The mean HCC risk scores between reported patients and non-reported patients were not significantly different. Therefore, no clear pattern of HCC risk differences between reported and non-reported patients appears to be emerging.
- Finally, we examined the degree of “adherence” of PQRI-reported beneficiaries to their PQRI reporting physician. In this study, we define “adherence” as the percent of all non-hospital E&M visits that the beneficiary obtained from a specific physician. Ideally, an “accountable” physician would be the physician providing most of the care to the beneficiary, e.g., exhibiting a high degree of adherence. At this point, however, there are no standard metrics for what “most of the care” means. A practical (empirical) upper bound on adherence is the percent of non-hospital E&M visits that the same set of beneficiaries obtained from the physician (reporting or not) *from whom the beneficiary obtained most of their visits*. We would expect a slightly lower percent for PQRI-reported patients because, if they had a PQRI-reportable condition, they are likely to be sicker and possibly receiving care from a greater number of physicians. That upper bound was 50 percent for PQRI-reported patients (52 percent for non-reported patients of the same physicians); whereas PQRI-reported patients had 34 percent of all their non-hospital E&M visits from the physician who reported their PQRI measure. (Non-reported patients of the same physicians received 36 percent of their visits from the PQRI reporting physician.) Thus, the percent of non-hospital E&M visits obtained from the PQRI reporting physician averages about 70 percent of the visits obtained from the physician from whom the beneficiary received most of their non-hospital E&M visits. The discrepancy between the upper bound and the level of adherence for PQRI-reported beneficiaries diminishes when the beneficiary’s health status is controlled.

These findings suggest that even though initial PQRI reporting rates are low, PQRI-reported beneficiaries are remarkably similar to non-reported beneficiaries seeing the same physicians, given that the PQRI-reported beneficiaries must have a reportable health condition. Indeed, we find that, even within broad diagnostic categories, PQRI-reported beneficiaries have more visits. Further, and most importantly, the levels of adherence for PQRI-reported beneficiaries are within 70 percent of the empirical upper bound, given the patterns of care exhibited by FFS Medicare beneficiaries in Colorado in 2008. Thus, to the extent that adherence is related to accountability, the PQRI measures appear promising.

B. Future directions in the development of the beneficiary level cost of care measure

In order to reward physicians for efficient and high “quality” care – one of the fundamental points of this entire effort – there must be some way to attribute costs to physicians that are in an ordinary sense “accountable” for those costs. But the question of accountability is

not entirely resolved by our analyses to date. Those analyses suggest that the physician who reports a PQRI condition for a patient is not necessarily the physician who most frequently treats that patient, but we do not yet know what kind of physician (i.e., by specialty code) is most frequently seen if it is not the PQRI-reporting physician.

If the validity of the attribution measure is dependent on the quantity of care the beneficiary receives from a physician, then by definition, using the post-hoc attribution rule “the physician from whom the beneficiary receives most of his/her care” will be the best measure. However, if one places greater emphasis on the physician’s willingness to accept responsibility for the care of the beneficiary, then a measure based on PQRI reporting, while not fully *ex ante*, at least has the advantage of involving active agreement on the physician’s part.

V. Next Steps

The next steps of our project involve five basic activities:

1. Further refinement of MS-DRG-based and beneficiary level attribution measures

Our work to date has made substantial progress on both types of attribution measures, but there is clearly additional work to be done. First, our work suggests that the MS-DRG approach shares the common problem with other episode grouper approaches of low numbers of cases for less common conditions or, in this case, events such as hospitalizations. We are not far enough along to determine how aggregation to TINs or how aggregation across MS-DRGs by physician may increase the number of observations. We are encouraged, however, as we explore a new approach for risk adjustment and other innovations that MS-DRGs can provide a credible means of holding physicians accountable for Part B expenses in closer alignment to efforts to hold hospitals accountable for Part A costs beyond the DRG payment. While it is too early to draw conclusions, our work to date seems to support the potential to include specialists such as surgeons in a resource use performance measurement system and to allow CMS to advance an alignment of inpatient related cost of care in non-integrated care settings.

Second, the beneficiary level analysis has produced new information on PQRI reporting that, to our knowledge, has not been studied prior to this work. These findings will be reported in the future. We continue to consider the meaning of the self attribution feature of PQRI reporting as a foundation for cost of care attribution, especially under attribution situations where two or more ambulatory care physicians have provided similar levels of care. However, we have found that beneficiaries with PQRI measures reported by physicians in the ambulatory setting do *not* receive a greater concentration of their care from the PQRI reporting physicians compared to similar patients (e.g. diabetics) with no PQRI measure reported. We will continue to use PQRI reporting combined with other means of improving the legitimacy of attribution, such as new risk adjustment methods and removing primary care non-attributable costs from a beneficiary’s annual cost calculation, to develop alternative attribution approaches for review by the Technical Expert Panel and CMS.

2. Clinical Performance Measures Development and Testing

Given the attribution measures developed and tested above for resource use, we seek to explore clinical performance measures that provide the other side of the value purchasing equation: the quality side. We are currently developing several approaches to creating evidence-based quality (and some overuse) performance measures. Each will build heavily on work done by others. We will rely on two levels of claims data for these measures:

- **Administrative data**, which is available on all encounters, and
- **PQRI data**, which is limited to those practices that voluntarily provide it.

We will review existing guidelines to ascertain which guidelines employ exclusively administrative data and which require more extensive clinical information. We will then create several levels of *performance* measures: 1) episode-specific measures; and 2) condition-specific measures. Episode-specific measures will focus on measures that are directly related to the episode in question, whereas condition-specific measures may include experiences involving multiple episodes or conditions not addressed by episodes. An example of an episode-specific measure would be a quality measure like use of beta-blockers or aspirin for AMI applied to an AMI episode. A set of condition-specific measures could be derived from various episodes like diabetes, AMI, congestive heart failure and asthma. In addition, some measures may not apply uniquely to specific conditions. For example, these measures may include structural elements such as record keeping or even staffing levels that are believed to improve quality across conditions.

In some instances, the performance measures will overlap well with the episode measures, but in other cases, they will not. For example, we may create a quality profile for a physician based on his/her performance across a number of areas, some of which are addressed in our episodes and some not. The “efficiency” score might then draw on data from two distinct areas of practice. Conversely, we may be able for some episodes to develop direct markers of good and bad quality.

A major task will be developing methods to combine measures to create a scale or score. We expect to work with a Technical Expert Panel (TEP) that will be engaged in spring 2010 to have them use techniques like magnitude estimation to generate weights for the various quality components.

3. Combine Cost and Clinical Performance Measures

In the final step of measure development, we will generate a set of cost measures and performance measures at the physician level for both the MS-DRG and the beneficiary level approaches. These will be developed in such a way as to allow CMS to use them individually, but we will also propose ways in which they could be combined to produce what amounts to a measure of efficiency that reflects the relationship between cost and performance. The goal is to identify physicians (or practices) that are generating the highest levels of performance and the lowest attributed costs. We propose two basic approaches to achieve the same end. The first is a

simple search algorithm to define a set of practices on the price-performance “frontier.” The performance of a physician practice is then computed as its distance from the high quality – low cost best performers. The second constructs “quality tiers.” A number of quality tiers are constructed using the distribution of performance on quality measures. The cost of care measures are then compared to practices within the same quality tier. The frontier analysis will ideally identify those practitioners who are delivering better quality at lower cost, or at those practices that are achieving one goal without endangering the second. The tier approach offers an alternative way to achieve the same effect by ranking practices by their quality measures and adjusting their within tier costs according to their quality scores.

4. Evaluate the fairness, meaningfulness, and potential to guide improvement of the measures

We will work with the clinical TEPs to evaluate the performance measures we have developed. The panels will review the measure prototypes that have been developed and tested and will consider options for attribution and peer grouping. The input from the panel will be used to refine the measures for evaluation on a nationally representative data set.

5. Reproduce measures on a representative regional sample of the US physician practices

In order to evaluate regional differences, the final step in the project will be to compute the measures across representative regions of the U.S.

VI. Conclusion

Any effort to improve the efficiency and quality of health care services in the U.S. will require physicians to take greater responsibility for their patients’ care. That step, in turn, depends not only on improved measures of efficiency and quality, but also on improved attribution methods that link beneficiaries to physicians.

In this project, we are using MS-DRGs as the basis for measuring and attributing inpatient care to physicians that provide primarily inpatient care, and we are exploring CMS’s PQRI reporting system as a basis for attributing beneficiaries to physicians who provide primarily outpatient care. Our analyses based on data from one year in one state are encouraging, although the PQRI system is so new that it is difficult to draw any firm conclusions.

In addition to exploring different methods of attributing inpatient surgical care (operating versus performing physician) in the MS-DRG analysis, we have developed and tested ways to group both inpatient and follow-up care (including both institutional and non-institutional post-acute as well as readmissions) and tested various lengths of observation “windows” associated with an initial admission, settling on 30- and 60- day windows. We have also begun development of risk-adjustment models for the MS-DRG analyses.

Our analyses of the PQRI reporting system found low rates of reporting by physicians

and low rates of beneficiaries for whom a non-hospital, E&M visit was reported. However, we found little evidence of purposeful selection or “cherry-picking” of patients for PQRI reports. We also found that the percent of non-hospital E&M visits obtained from the PQRI-reporting physician was approximately 70 percent of the upper bound defined by the physician from whom beneficiaries received most of their non-hospital E&M visits.

During the next reporting period, we will continue our work on both the MS-DRG and beneficiary level analyses. We will also begin our development and testing of quality measures, further risk-adjustment methodologies, and combined measures of cost and quality.