CLINICIAN FEEDBACK ON USING EPISODE GROUPERS WITH MEDICARE CLAIMS DATA

January, 2010

Centers for Medicare & Medicaid Services
Office of Research, Development, and Information

Todd Caldis, Ph.D. (Office of the Actuary)
Craig Caplan, M.A. (ORDI)
Marty Cohen, MPA (Kennell & Associates)
James Leonard, MPH (Kennell & Associates)
Jesse Levy, Ph.D. (ORDI)
Curt Mueller, Ph.D. (ORDI)
Fred Thomas, Ph.D. (ORDI)

TABLE OF CONTENTS

TABLES

CLINICIAN FEEDBACK ON USING EPISODE GROUPERS WITH MEDICARE CLAIMS DATA

Executive Summary

The Centers for Medicare & Medicaid Services (CMS) is investigating techniques that might help identify costly physician practice patterns. One method presently under evaluation is to compare resource use for certain episodes of care using commercially available episode grouping software. Although this software has been used by the private sector to classify insured individuals' medical claims into episodes of care, it has never been used with fee-for-service Medicare claims except in the studies by the Medicare Payment Advisory Commission (MedPAC) and CMS.

Episode groupers are proprietary software programs that organize claims data into a set of clinically coherent episodes, usually linked by diagnosis. Two proprietary episode groupers were available: INGENIX Symmetry's Episode Treatment Groups (ETG) and the Thomson / Reuters Medical Episode Grouper (MEG). The episode groupers build episodes of care using all of a beneficiary's health care claims that are filed between two points in time for a specific health problem. The aggregate dollar amount of claims payment constitutes the cost of an episode. Thus, the design issues of what claims to include in an episode and how to include them are relevant to studying an efficient level of cost for an episode type. Conceptually, episode costs can be attributed to a physician and then compared across physician practices.[1]

[1] Ideally, physician efficiency should also include outcome measures or other measures of quality. Low cost, by itself, is not necessarily a desirable goal.

The episode grouping software requires users to specify the input parameters for a given set of outputs. The user determines the types of claims data that will be grouped, the time frames for which the data are collected, the various software profile settings, the physician attribution and benchmarking algorithms, and the outputs that will be provided to clinicians, among other decisions. Because of this flexibility and the substantial user involvement, no one correct episode grouping method exists.

To construct an episode, claims are generally linked by diagnosis code(s). Episode types can be characterized as chronic, acute, or preventive care. Chronic condition episodes are defined for episode grouping purposes as having a 12-month duration, usually a calendar year, even though by definition, a chronic condition does not end. Acute and preventive care episodes can start at any time during a year and may continue into a subsequent calendar year.

Eight panels, which in aggregate consisted of approximately 80 clinicians, were asked to respond to a series of episode grouping issues, which were identified after studying Medicare claims data that had been grouped with two commercially available software products. This study reviews and reports on clinician feedback on the most obvious and important decisions that must be faced by Medicare to use grouped claims data as the foundation for a physician performance measurement system. Specifically, eight episode grouper design issues were discussed with the panels: (1) grouping physician claims with an inpatient hospital stay, (2) grouping an inpatient stay with SNF claims, (3) grouping an inpatient stay with home health claims, (4) excluding certain claims types, (5) grouping complications of medical and surgical

care, (6) grouping acute exacerbations of a chronic condition, (7) grouping signs and symptoms (non-specific) diagnosis codes, and (8) defining the duration of chronic episodes.

Representative of the discussions is a relatively straight forward case: an episode that involves a hospitalization. The issue is if physician professional fees that are rendered during the hospitalization should be grouped into the same episode as the hospital claim. The grouping of inpatient Part B physician claims is complicated because Medicare patients typically have co-morbid conditions that also are treated during an inpatient stay. Consequently, inpatient physician claims associated with the treatment of conditions in their specialty may include a diagnosis code that differs from the principal diagnosis shown on an inpatient facility claim. Evidence of differences in diagnosis was found by MaCurdy (2008) who found a high percentage (44 percent in ETG; 60 percent in MEG) of the Part B physician claims provided during an inpatient stay were assigned to an episode that did not include the inpatient hospital facility claim when using diagnosis for grouping.

Since Medicare's payment for a physician service is based on the CPT-4 code (reflects procedure or type of visit) rather than on the diagnosis, physician offices have no incentive to spend much effort in coding a diagnosis. In contrast, the payments hospitals receive are determined by a combination of diagnosis and procedure codes. Thus, the diagnoses shown on inpatient physician claims may not be as accurate as they are for inpatient hospital services.

An alternative available in one grouper is to group all claims (e.g., professional Part B, SNF, outpatient, home health) that occur between the admission and discharge dates of an

inpatient stay into the same episode as the inpatient stay. While this method will group all professional Part B claims with the associated inpatient stay episode, the services may have little to do with the principal diagnosis assigned to the inpatient claim. For example, a psychiatric consult will have a different diagnosis than congestive heart failure and may have little to do with the reason for the hospitalization.

The panels also raised concerns about validating the grouper logic, risk adjustment, homogeneity of episode costs, adequate sample size, the validity of peer groups, transparency, actionable information, quality performance, and rural issues. The panel reactions show the importance of bringing persons with clinical knowledge into the development process and that additional research is needed.

Introduction

Policy makers have expressed concerns that the current Medicare payment system includes incentives that encourage physicians to overuse some services and underuse others; pays physicians for care irrespective of their level of resource use; and offers higher revenues to physicians who furnish more services, regardless of whether they add value (MedPAC 2008). To remedy these concerns, the concept of value based purchasing (VBP) has been introduced into the policy arena. The goal of a VBP program is to find ways to reward physicians financially for providing efficient use of resources and services that are of high quality. The Medicare Improvements for Patients and Providers Act of 2008 (MIPPA) requires the Department of Health and Human Services (DHHS) to develop a plan that will transition Medicare payments into a VBP program for physician and other professional services that is

based on efficiency and the quality of services provided. The Act also requires the DHHS to disseminate informational reports to physicians using episode groupers and/or per capita measures.

CMS has been investigating techniques that can help identify higher cost practice patterns. One technique is to compare resource use at the episode of care level. Episodes of care represent a group of healthcare services (claims) for a health condition (e.g., hip fracture, diabetes) over a defined length of time for which a physician can be responsible. The Medicare Payment Advisory Commission (MedPAC) has argued that compared with traditional population-based metrics, episode measurement could:

- Allow the attribution of care to individual physicians;
- Avoid focusing on a narrow set of utilization measures at the expense of others;
- Help identify specific changes in practice that can improve cost efficiency relative to peers; and
- Provide better accounting for differences in inpatient health status (MedPAC 2005).

Existing "Episode groupers" are proprietary software programs that organize an individual's claims into clinical episodes of care. In its March 2007 report, MedPAC stated that "episode groupers can be used with Medicare data." After two years of study with Medicare claims data, MedPAC concluded that episode groupers have face validity from a clinical perspective, can identify practice patterns, and have risk adjustment capabilities that can account for differences in disease severity and the presence of co-morbidities (MedPAC 2007).

While existing episode groupers have been developed for and used in private sector health plans and insurers as management tools, these grouping software programs have never been used in fee-for-service Medicare. As a result, Acumen, LLC under contract to CMS began studying episode groupers using 2006 Medicare claims data. Acumen produced two publicly available reports: one on the functionality of the two commercially available groupers, and the other on issues in using grouped episodes to create resource utilization reports.[2] These studies use quantitative analysis to explore how the episode groupers work with Medicare claims data.

This study focuses on the major design issues that should be addressed if episode grouping software is to be used with Medicare claims data. Eight panels, which in aggregate consisted of approximately 80 clinicians, were asked to respond to a series of episode grouping issues, which were identified after studying Medicare claims data that had been grouped with two commercially available software products. The panels were not asked to determine the most appropriate commercial episode grouper for Medicare claims, but to discuss design issues that must be considered in developing episodes in a value based purchasing context regardless of the episode grouping software used.

Background

Episode groupers are proprietary software programs that organize claims data into a set of clinically coherent episodes, usually linked by diagnosis. Two proprietary episode groupers were available: INGENIX Symmetry's Episode Treatment Groups (ETG) and the Thomson /

[2] See http://www.cms.hhs.gov/Reports/downloads/MaCurdy.pdf and http://www.cms.hhs.gov/reports/downloads/MaCurdy2.pdf

Reuters Medical Episode Grouper (MEG). The episode groupers build episodes of care using all of a beneficiary's health care claims that are filed between two points in time for a specific health problem. The aggregate dollar amount of claims payment constitutes the cost of an episode. Thus, the design issues of what claims to include in an episode and how to include them are relevant to studying an efficient level of cost for an episode type. Conceptually, episode costs can be attributed to a physician and then compared across physician practices.[3]

The episode grouping software requires users to specify the input parameters for a given set of outputs. The user determines the types of claims data that will be grouped, the time frames for which the data are collected, the various software profile settings, the physician attribution and benchmarking algorithms, and the outputs that will be provided to clinicians, among other decisions. Because of this flexibility and the substantial user involvement, no one correct episode grouping method exists.

To construct an episode, claims are generally linked by diagnosis code(s). Episode types can be characterized as chronic, acute, or preventive care. Chronic condition episodes are defined for episode grouping purposes as having a 12-month duration, usually a calendar year, even though by definition, a chronic condition does not end. Acute episodes can start at any time during a year and may continue into a subsequent calendar year. Acute and preventive care episodes are generally much shorter in duration, and are considered complete or "closed" when there is no activity within a given period, i.e., when a "clean period" is reached. For example, if no claims for hip fracture are encountered 90 days after the last such claim, and the clean period

[3] Ideally, physician efficiency should also include outcome measures or other measures of quality. Low cost, by itself, is not necessarily a desirable goal.

is defined to be 90 days, then the episode is deemed complete. Only certain claim types can initiate or open an episode. In both episode groupers, physician Part B evaluation and management claims or surgery claims, an inpatient hospitalization, or a skilled nursing facility stay can start an episode. Ancillary claims and durable medical equipment claims cannot start an episode in either episode grouper, while home health claims and hospice claims may start an episode in one episode grouper. Since a person can be treated for many conditions at the same time, episodes may be open simultaneously for different conditions.

Few independent evaluations of episode grouping software have been published, and only two studies, both funded by CMS, have used Medicare claims data. The first CMS report on episode grouping by Acumen, LLC, entitled "Evaluating the Functionality of the Symmetry ETG and Medstat MEG Software in Forming Episodes of Care Using Medicare Data," concludes that 1) the grouping algorithms do not emulate practice patterns common in the Medicare system, 2) inpatient physician services often do not group with the associated hospital stays, and 3) there is considerable variation in costs across episodes and within episode types." This large variation in episode costs "…suggests the need to develop models of risk or severity adjustment applicable for Medicare populations prior to being able to use the episodes … software for profiling Medicare providers" (MaCurdy, 2008-A).

A second CMS report by Acumen, entitled "Prototype Medicare Resource Utilization Report Based on Episode Groupers," discusses practical aspects of implementing resource utilization reports using episode groupers. Options are given for apportioning cost into episodes, assigning physician attribution rules, defining relevant peer groups, adjusting costs, and

developing peer group cost benchmarks. Empirical analysis is performed on the number of episodes required for valid comparisons (MaCurdy, 2008-B).

Methods

In the summer of 2008, two panel discussions were organized at each of four large multi-specialty group practices: Geisinger Health Systems in Danville, Pennsylvania; Billings Clinic in Billings, Montana; St. John's Health System in Springfield, Missouri; and Marshfield Clinic in Marshfield, Wisconsin and Wausau, Wisconsin, for a total of eight panels. For each panel, 8 to 12 clinicians met for approximately three hours to discuss episode grouping issues, a total of 80 mostly physician participants for all of the panels. The composition of each panel was influenced by the type of condition that was to be discussed. For example, orthopedic surgeons were overrepresented on a panel that focused on hip fractures. Most panel participants were not familiar with episode grouping software and value based purchasing concepts. One site, however, had used an earlier version of one episode grouper to monitor resource use in their health plan division; although clinicians at this site were aware of the grouping basics, they were not familiar with the discussion questions and design issues that would be posed.

After listening to a brief presentation on value based purchasing, episode grouping basics, and episode design issues, the panel members were asked to discuss the types and combinations of health care claims for which a physician should be held accountable. In addition to examining overarching episode design issues, each panel also focused on one of the following five clinical conditions: chronic obstructive pulmonary disease (COPD), hip fractures, diabetes, congestive

heart failure (CHF), and coronary artery disease (CAD).[4] Two clinical conditions were discussed at each site. A physician under contract to CMS made the introductory presentation to each panel and led the subsequent discussions. Kennell & Associates, Inc., under contract to CMS, organized the panels, conducted the meetings, and prepared notes. CMS staff also attended the meetings.

The presentation on value based purchasing and episode grouping basics was prepared by CMS and Kennell & Associates, Inc. staff using policy papers and software documentation provided by the vendors of the commercial episode grouping software that were used to prepare claims data for review by the panels. Episode design issues were developed in advance of the panels by reviewing the issues identified in Acumen's first CMS report on episode grouping (MaCurdy, 2008) and by examining how the two episode groupers linked Medicare claim types at each site. Using the algorithms developed for the 2008 report on functionality, Acumen, LLC then grouped the 2004-2006 claims at each study site, so that each panel would be exposed to examples of how their own claims had been grouped. Design issues were identified by comparing the relationship of grouped claims and claims types across episodes to expected treatment patterns. For example, how were inpatient physician claims grouped when the physician claim diagnosis differed from the inpatient hospital diagnosis? This paper identifies and explains significant design issues, and then discusses the reactions to them by the clinician panels.

[4] These conditions were selected from a subset of acute and chronic conditions that are highly prevalent in the Medicare population.

Claims analysis was conducted at the "base episode" level, without incorporating risk adjustment methodologies or severity levels (i.e., 4 disease stages in MEG and up to 4 severity levels in ETG). In this way, the panels could more easily focus on the fundamental design issues. If a discussion required an explanation of these adjustment methodologies, the panel was given this information. Rather than quantifying positions taken (such as the percentage advocating a particular position), this study used a qualitative approach to synthesize the positions, arguments, and insights provided by the approximately 80 clinicians who voluntarily participated. Summarized below are the panel discussions, which reflect the thoughts of the panels, and not CMS.

Summary of Clinician Feedback

The discussions were characterized by the following design issue areas:

- Grouping of claim types: whether and how certain types of claims, such as inpatient hospital (IP), physician, skilled nursing facility (SNF), and home health (HH) claims, were grouped into episodes;
- Related clinical events: how related clinical events, such as acute exacerbations of chronic conditions and complications of surgery, should be grouped.
- Other issues: such as how the length of chronic episodes should be defined and what claims types should be excluded.

Discussions on the three issue areas are summarized below using the following format: 1) identifying the grouping issue(s), 2) summarizing the major thinking and/or consensus reached by the panels, and 3) listing the discussion points made at each site in detailed table. In addition, overarching concerns about using episodes to measure physician performance were also expressed at each site. While most of these concerns were not the subject of the formal

presentations, the panel members believed they were important. Consequently, these concerns and feedback are included in this report and summarized in the last section.

Grouping of Claim Types

Grouping Physician Claims with an Inpatient Hospital (IP) Stay

During an inpatient hospital stay, patients often receive services from many physicians. The grouping of inpatient Part B physician claims is complicated because Medicare patients typically have co-morbid conditions that also are treated during an inpatient stay. Consequently, inpatient physician claims associated with the treatment of conditions in their specialty may include a diagnosis code that differs from the principal diagnosis shown on an inpatient facility claim. Since the diagnosis code is normally used to link claims into one episode, matching Part B physician claims with an inpatient facility claim may be problematic if the diagnosis codes differ.

As an example, a patient hospitalized for a hip fracture may also need to be treated by an internist for hypertension. In this case, the Part B claim from the internist's evaluation and management service could have a diagnosis code for hypertension, but not a diagnosis (or procedure code) for hip fracture. In contrast, the inpatient hospital claim would show a principal diagnosis related to the hip fracture procedure (as would the diagnosis on the surgeon's Part B claim). The episode grouping software must assign the internist's Part B claim to one and only one episode: that is, either to the hip fracture episode or to a hypertension episode. Since

Medicare's payment for a physician service is based on the CPT code (reflects procedure or type of visit) rather than on the diagnosis, physician offices have no incentive to spend much effort in coding a diagnosis. In contrast, the payments hospitals receive are determined by a combination of diagnosis and procedure codes.

Based on grouping options available in one episode grouper, two design strategies can be used to link inpatient facility and professional Part B claims. In this study, these strategies are termed the "diagnosis method" and the "date of service method." The "diagnosis method" assigns Part B claims to episodes during an inpatient stay according to the diagnosis (and sometimes procedure) on the claim. In linking claims primarily by diagnosis, a high percentage (44 percent in ETG; 60 percent in MEG) of the Part B physician claims provided during an inpatient stay was assigned to an episode that did not include the inpatient hospital facility claim (MaCurdy, 2008). In contrast, the "date of service" method assigns Part B claims during an IP hospital stay according to the date on the Part B claim, regardless of the diagnosis information on that claim. This option will also group claims for unrelated conditions into that episode. The key points from the panel discussions shown in Table 1 are summarized below.

Use the date of service method: Some physicians concluded that much of the care would not have been done without a hospitalization; therefore, according to this perspective, all physician services during an inpatient stay should be grouped into the same episode as the inpatient stay. The diagnoses shown on inpatient physician claims may not be accurate, since Medicare physician payments are not determined by diagnoses, as they are for inpatient hospital services. Therefore, using inaccurate diagnoses to link claims may not construct reliable episodes.

Use the diagnosis method: The prevalence of chronic co-morbidities often requires multiple inpatient consults for different conditions; therefore, some panel members argued that physician claims should be grouped by diagnosis. A single patient may have multiple concurrent episodes of illness. Grouping all inpatient consults into the same episode as the inpatient stay may provide an incentive for physicians to avoid treating co-morbid conditions during an inpatient stay. Therefore, some panelists favored using the diagnosis on the Part B claims for grouping.

Table 1: Feedback on Grouping Physician Claims With an Inpatient Hospital Stay

Site 1 (Names blinded)

The "date of service method" should only include claims for related diagnoses. It might make sense to include professional fee claims for other diagnoses within the inpatient stay if the treatment was for a diagnosis related to the inpatient diagnosis or a co-morbid condition that impacts the inpatient diagnosis. Part B claims for diagnoses that are not related and that do not impact management of the inpatient diagnosis should be excluded.
The "date of service method" depends on the reliability of the risk / severity adjustment. If the severity risk adjustment were done well, then physicians would be more comfortable including more of the Part B professional fee claims that occur during an inpatient stay with the inpatient stay. If risk adjustment is not accurate, it is more important to split out claims for other diagnoses during the inpatient stay.
The "date of service" method could result in fewer physicians meeting the threshold for getting a Resource Use Report. If all claims that occur during an inpatient stay are bundled together, each physician will have fewer episodes, and this might reduce the number of physicians who have enough episodes to be included in the efficiency reports. It also might decrease the ability to risk adjust.
Diagnosis codes on Part B claims might not be accurate. Physicians believe that the diagnoses listed on the Part B professional fee claims that occur during an inpatient stay might not be accurate since reimbursement issues might in some circumstances be driving coding use (concurrent care).

Site 2

Most of the panel members were in favor of using the “date of service method” to group the physician claims during an inpatient stay, but ideally unrelated claims should be excluded. Clinicians generally expressed the opinion that to measure resource use, the episode should group inpatient care to the diagnosis that caused the exacerbation and the hospitalization. Even when the patient is seen by many physicians, much of the care during the stay would not have occurred if the patient had not been hospitalized for that condition. While it would be good to exclude Part B claims for diagnoses that are completely unrelated, it would be difficult to figure that out without reviewing the medical record.

The “date of service” method could lead the system to undervalue the cost of caring for chronic conditions. If this method is used, the inpatient care for chronic conditions that are treated during a hospitalization for an acute condition will be attributed to the diagnosis for which the patient is admitted rather than to the chronic condition.

Site 3

Most panel members favored using the “date of service.” Clinicians generally expressed the opinion that to measure resource use, the grouper should group all inpatient care to the diagnosis that caused the hospitalization. While it would be beneficial to exclude Part B claims for diagnoses that are completely unrelated, it would be difficult to determine this without reviewing the medical record. It is difficult to know where to draw the line on whether inpatient professional care is or is not related to the cause of the hospitalization.

The orthopedic surgeons said surgeons, in general, do not encourage "a lot of extraneous evaluation" of a patient, so most inpatient care for a surgical patient should be related to the surgical episode and therefore using “date of service method” should be reasonable.

There was concern on whether sample sizes would be large enough, which will depend in part on what rules and methodologies are adopted for reporting. Assuming a large enough sample size and adequate risk adjustment, physicians seemed comfortable that the impact of unrelated inpatient care could average out when compared to the peer group.

One panel member was concerned that the “date of service” method would provide an incentive not to treat co-morbidities during an inpatient stay, reducing overall quality of care. One physician was concerned that if the cost of all inpatient consults were attributed to one responsible physician, then that physician would have an incentive to minimize treatment of co-morbid conditions during an inpatient stay. This would make that physician look more efficient, but could reduce quality of care if it meant patients were not receiving timely care for co-morbidities.

Site 4

Several panel members thought the “date of service” method would provide the wrong incentives to physicians. These panel members were concerned that if this method were

used, physicians would have a financial incentive to avoid treating co-morbidities during a hospitalization, because they would be worried about picking up these additional treatment costs as part of the basic inpatient episode. One physician noted that because they serve patients in a rural area, patients who were hospitalized were often treated for multiple problems. Thus, in rural areas, "date of service method" might not be as accurate as in urban areas.

Other panel members thought neither approach was clearly superior to the other, and that the key principle would be to measure all physicians using the same method. These panel members thought grouping errors would be unavoidable under either approach, and that we don't know which method would make fewer errors. As one physician noted, "There is no magic answer." Therefore, the key principle would be to measure all physicians using the same method, so that, in theory, the same degree of error would also be reflected in the peer group.

Some panel members wondered whether splitting out different diagnoses would encourage hospitalists (and other physicians) to start using different diagnoses for multiple visits with the same patient.

Another physician was concerned that using the "date of service" method could send a message that the specificity of a physician's diagnosis coding for inpatient claims is not important. The physician was concerned that if physicians knew that all their inpatient visits would be grouped based on the hospital claim's diagnosis, regardless of diagnosis on their own individual claims, then physicians would pay less attention to how they code diagnoses on their claims (at least for inpatient services).

Outpatient Facility (OP) Claims

Outpatient (OP) facility services include emergency room (ER) visits, laboratory, radiology, and ambulatory surgery. Since OP claims are billed on the same billing form as inpatient hospital (CMS 1450), multiple services and multiple diagnoses can be recorded on one claim. Matching one of the multiple diagnoses with the correct line item on the claim may present a problem in grouping (the first diagnosis is assumed to be the principal), although some grouping logic may help in this regard (particularly in identifying ancillary services). Unlike Part B professional bills, there is no reference between the header diagnosis and individual line

items on Part B OP claims. (On a Part B claim, each service has one and only one diagnosis assigned to it from the list of diagnoses on the header of the claim, while on an OP claim, a single diagnosis for a service is not provided.)

Medicare considers an outpatient facility charge as part of the inpatient hospital payment if the outpatient facility claim occurs on the same day or precedes the admission by 72 hours. This Medicare payment rule may complicate the grouping of ER and other OP facility services, since separate claims may not be generated if they are related to an admission (the cost or payment will be reflected in the inpatient claim, and will not result in an OP claim). Without an inpatient admission, the ER and/or OP services would result in separate claims that would be grouped to an episode. Even if identical services are furnished, in one case, the ER and/or OP services claims would increase the costs of an episode, while in the other case, the same claims would not generate any new costs of an episode. There will also be professional Part B claims that may or may not have a corresponding facility charge in the episode, depending on the admissions status.

Both groupers rely almost entirely on the principal diagnoses to link claims in an episode, although other attributes of a claim may be used to "break ties" between assigning a claim to one episode over another. There is no separate logic that links associated or sequential services, such as grouping an ER visit and an accompanying inpatient stay, an anesthesiology claim with the accompanying surgical procedure, or an ER facility claim with the corresponding professional Part B claim(s).

The questions posed to the panels relate to what diagnosis information to use and how the inpatient bundling of OP services should be reflected in an episode:

- Should only the first diagnosis be used to group these claims, even if the claim has multiple diagnoses?
- Should different OP claims on the same date automatically be grouped together, even if they have different diagnosis codes?
- Should an OP claim preceding an inpatient stay within a certain time window automatically be grouped to the same episode as the inpatient stay?

Table 2 lists the specific comments from the panel discussions. In general, panel members favored reliance on the principal diagnosis on the OP claim to group OP claims, rather than considering secondary diagnoses or date proximity to another OP claim or to a hospitalization. Although a few physicians favored grouping all OP claims within a certain time window together, most thought that the diagnosis codes on the claims should be used to guide the grouping decision for each claim, in which case the different OP claims on the same date would not be grouped together automatically. Almost all the physicians thought it would be inappropriate to group automatically all OP claims within a certain time window of a hospitalization as part of the inpatient episode.

Table 2: Feedback on Outpatient Facility (OP) Claims

Site 1 (Names blinded)

The primary diagnosis on the outpatient claim is initially assigned by the physician. The laboratory may change the primary diagnosis on the outpatient claim. If the laboratory work is done at a community hospital instead of at the practice, a technician at the community hospital might change the order because they do not know the practice's system. Also, the lab is focused on finding the best way to get paid and will try to pick a diagnosis with that in mind. Ultimately, diagnoses listed on outpatient claims may not be reliable.

Site 2

The primary diagnosis on the outpatient claim is initially assigned by the physician.

Site 3

It would be preferable to rely on the primary diagnosis for these claims. There was consensus that as long as an outpatient facility claim has a specific diagnosis as the primary diagnosis, then the grouper should assign the claim based on that diagnosis, rather than relying on secondary diagnoses, information on procedures, or some type of "outpatient build" approach (based on proximity to dates of service).

Site 4

Many patients get different lab tests at the same time, for multiple conditions. A patient often might see two physicians who both order lab tests, in which case the physicians often coordinate their orders so that the patient gets all the lab tests in one appointment. This is more efficient for the patient (patient only has to fast once, only schedules one appointment, etc.), but this means lab claims often might include multiple tests that truly were for different conditions and different episodes of care. The primary diagnosis on the lab claim may not reflect the true primary diagnosis. The practice has electronic ordering of lab tests, but the order of the diagnoses generated for a lab claim often reflects the alphabetical order of the diagnoses or of the tests being done, because often the physicians chose the relevant tests or diagnoses in the order they are presented on the screen. One physician said she was not aware of the sequence for tests or diagnoses on the electronic order would matter for anything, and other panel members suggested this was probably a common view. This suggests that the primary diagnosis on lab claims with multiple diagnoses may not be very reliable, even for group practices with electronic ordering. Neither panel liked automatically grouping OP claims preceding an inpatient stay within a certain time window. Members of both panels said they have many cases of patients getting multiple ancillary services for different conditions on the same day, so the physicians would be uncomfortable with a grouping rule that automatically assigned all these claims to the same episode based on date proximity, regardless of their diagnosis codes. For outpatient services that precede an inpatient stay, the physicians also were uncomfortable automatically linking these outpatient claims to the inpatient episode regardless of diagnosis, because patients might be seen for a routine outpatient visit and then something unrelated is discovered which leads to a hospitalization. Under this scenario, the physicians did not believe the routine outpatient visit should automatically be grouped with the inpatient episode, if the diagnoses were unrelated. For example, a patient might get a routine physical or mammogram, and then a more serious condition is discovered which ultimately requires hospitalization. The physicians believed the routine physical or mammogram should not be grouped with the inpatient episode.

Grouping an IP Facility Claim with Skilled Nursing Facility (SNF) Claims

Medicare pays for post-acute care in skilled nursing facilities (SNFs) following an inpatient stay (having duration of at least 3 days) and related to the condition of the hospitalization. Based solely on matching the principal diagnoses using the "diagnosis method," Acumen, LLC found that both groupers link SNF claims that immediately follow an IP stay (not necessarily the qualifying inpatient stay) in only about 50% of the episodes that contain the inpatient facility claim. However, the "date of service" method can be adapted for SNF claims in one episode grouper. The date of service method will group a SNF claim in the same episode as the IP claim only if the date of the hospital discharge is the same as the date of admission to the SNF. With this adaptation, 95 percent of SNF stays following an IP stay will be grouped into the same episode as the inpatient hospital stay. The panelists were asked about the appropriateness of applying the "date of service method" and the "diagnosis method" in grouping Medicare SNF claims immediately following a hospital stay. The discussion points are detailed in Table 3 and summarized below:

Use the date of service method: A SNF stay immediately following a hospitalization is usually related to the same medical condition as the inpatient stay. Including SNF claims in the same episode as the inpatient hospital claim may give physicians an incentive to manage the inpatient stay better in order to avoid a SNF placement.

Use the diagnosis method: The SNF stay may not necessarily relate to the preceding IP hospital stay. Some patients can have multiple SNF claims following an IP hospital stay. In these situations, the second SNF stay may reflects the patient's underlying condition or frailty rather than the acute condition that required hospitalization, making it most sensible to group the SNF claim separately from the episode containing the claim for the preceding inpatient stay. For example, dementia and other mental impairments listed as the triggering diagnosis on a SNF claim may indicate SNF care in order to ensure adequate post-acute treatment. In these cases the SNF cost should be attributed to a dementia episode, not the condition that caused the hospitalization. In these examples, if SNF care is always included with the inpatient episode the quality of care might suffer since physicians may have a financial incentive to avoid SNF placements.

There also were mixed opinions on whether to exclude SNF claims altogether from episode grouping. Some physicians noted that SNF claims should not be excluded, since SNF care is part of a continuum of care, and the whole continuum needs to be included when measuring physician resource use. Others argued that the variation in SNF usage is sometimes driven by factors that are beyond the physician's control (and which might not be adequately accounted for in risk adjustment or definition of a peer group), including the degree of local availability of SNF care or its substitutes (like home health) or the degree of support in the patient's home.

Table 3: Feedback on SNF Claims

Site 1 (Names blinded)

SNF diagnoses may not be accurate. SNFs know that a "deconditioned" diagnosis will get them a higher payment, so they tend to use these. For this reason, the SNF diagnosis was not considered reliable. Reasons for hospital and SNF stay may be different. In many cases, the SNF stay is not for the same reason as the immediately preceding inpatient stay. Therefore, the SNF should not automatically be assigned to the same episode as the inpatient care. Clinicians responsible for the hospital and SNF stays may be different. Medicine is increasingly using hospitalists for inpatient care. It is rare that the same clinician who manages the inpatient stay also manages the SNF stay. SNFs are also now using "SNFists." The use of SNFists or hospitalists should not significantly increase the costs of episodes unless this person causes serious complications. Including SNF in an episode may reward physicians who manage the hospital stay well. Some physicians are going to manage the inpatient stay better so that the patient does not require a SNF stay. In that case, the physician should get credit that efficiency. Combining claims for SNF and hospital care might be inappropriate in systems that do not use primary care managers or medical homes. Issues of attribution are a concern when combining claims for hospital and SNF stays. Instead of assigning 100% accountability to one physician, assign percentage weights of the care and split up a total episode payment. If the patient has a medical home, give the total episode payment to them and they would parcel it out. But operationally, that would create many challenges.

Site 2

In some cases, the SNF facility charge should be included in the same episode with the hospital stay. Even if the diagnoses are different, many SNF stays result from the patient becoming de-conditioned after the hospital stay (such as in COPD) or from the patient needing rehabilitation prior to returning home (as in the case of hip fracture). In these cases the SNF stay is part of the continuum of care. If a patient was in a SNF prior to the hospital stay and returns to the SNF after discharge, how much of the SNF stay can be attributed to the same condition that was the focus of the hospital stay? Part B claims during a SNF stay should not automatically be included in the same episode as the hospital stay. A SNF patient may be being treated for many conditions unrelated to the hospital stay; therefore, these costs should not be included with the hospital episode.

Even when the SNF costs are for the same condition as the hospitalization, determining which physician should be responsible is difficult. The physician managing the hospital care is not likely to be the same physician managing the SNF care. Particularly in rural settings, the physicians may be in different communities separated by hundreds of miles, with little influence over what happens in the remote setting.

Many factors over which the physician has no control influence the use of resources once the patient leaves the hospital. Resource use is heavily influenced by the willingness of the family to provide care and the availability of community resources such as outpatient, rehabilitation, and home care services which are more of a problem in rural areas.

Including SNF costs could confound measurement of physician efficiency. Physicians want to receive feedback and be measured on what they have control over. The farther out in time from the hospitalization, the more likely it is that confounding factors not under the control of the physician will be included. Introducing these confounding factors may cause the reporting system to miss the target of measuring physician efficiency.

Including SNF costs could lead to gaming and negative incentives for providing appropriate care. When something is measured, incentives are created and including SNF care could encourage a physician not to use it even when it is in the best interests of the patient. On the other hand, not using SNF care might make a re-hospitalization more likely.

Site 3

The panel members were divided on whether to include SNF costs when measuring a physician's episode costs and efficiency. One orthopedic surgeon commented that because SNF is so common after hip surgery, the SNF costs should be included in the surgeon's episode. However, another surgeon noted that whether or not a specific patient would need SNF care after surgery often depended on the degree of support available at home, which would have nothing to do with the surgeon's efficiency. The first surgeon's rejoinder was that as long as the sample was large enough, and as long as there is a risk adjustment to account for pre-existing conditions, then the risk for SNF costs should work out in the averages (if there were a sufficient number of patients). This same concern for a sufficient sample size and a risk adjustment was also raised by the cardiology panel when asked about including SNF costs in their episodes. One of the surgeons also acknowledged that he would probably pay a lot more attention to how long his patients were in the SNF, and why, if he knew the length of the stay would affect his financial bottom line.

The orthopedic surgeons differed from the cardiologists on whether to use the “date of service” method or the SNF diagnosis when assigning SNF claims to episodes. The surgeon’s view was that if a patient is in a SNF after hip surgery, it is reasonable to assume this to be the reason for the SNF admission. A cardiologist noted that many of his older patients suffer from dementia, and that often they might end up having a long SNF stay because of the dementia, and not because of their CAD. Therefore, the cardiologists' view

was that SNF should be assigned to episodes based on the SNF claim's diagnosis, not based on the "date of service" method linkage.[5]

Site 4

One concern raised was that if SNF costs are included when measuring physician efficiency, quality of care might suffer because physicians would have a financial incentive not to place patients in SNFs. However, two responses were provided for this concern. First, an overall efficiency measurement should include a quality component in addition to a cost component. Second, even within the cost component this incentive should be mitigated by the risk that skimping on SNF care could lead to complications and re-admissions, which could prove more costly than the SNF care.

Availability of SNF care varies widely by area, especially for rural areas. Variation in SNF costs among different physicians' episodes may be influenced more by SNF availability in different areas than by any aspect of physician efficiency. Physicians were very concerned that the availability of SNF care would not be adequately controlled for in the definition of the physician peer group.

Patient need for SNF care also is often driven by whether or not the patient has social support at home. This can vary significantly across patients and circumstances, but is not related to physician efficiency. This issue could be ameliorated if the peer group ranking reflected a similar mix of patients with and without social support in the home.

If SNF claims are excluded, claims for inpatient rehabilitation hospitals also might need to be excluded. Panel members noted that inpatient rehab hospitals can be substitutes for SNF care. If a decision is made to exclude SNF claims, then inpatient rehabilitation hospital claims would also need to be excluded to be consistent.

Condition-specific decisions on whether to include SNF care for efficiency measurement should be considered. CMS could decide to include SNF care when measuring physician efficiency for conditions with a high probability of SNF care, such as hip fracture, while excluding SNF care for conditions for which SNF care is less common, such as diabetes or CHF.

If SNF claims are to be included, their grouping should be based on the SNF diagnosis rather than use the "date of service" method (although this might not always be appropriate). In the CHF panel, there was a general consensus not to use "date of service" method with SNF claims, although there was recognition that grouping based on the SNF diagnosis could be inaccurate. In the diabetes panel, the opinions were mixed. One physician argued that if the SNF claim's primary diagnosis is completely different from the hospital's primary diagnosis, then the SNF claim should not be automatically linked to the hospital episode. Another physician countered that most of the time a SNF stay after a

[5] This would also be appropriate for patients who went to the hospital from a SNF and then returned to a SNF after the hospitalization.

hospitalization probably really is related to the hospitalization, whether or not the SNF claim is coded with the same diagnosis.

Grouping an IP Facility Claim with Home Health (HH) Claims

In contrast to SNF, a prior hospitalization is not needed to access the Medicare Home Health (HH) benefit. However, HH services immediately following a hospital stay can be grouped into the same episode as the inpatient stay using the date of service method, or be grouped by the diagnosis method. The need for automatically grouping HH claims into the same episode as the prior hospitalization was questioned by many panel members, who argued that HH should often be excluded from the same episode as the hospitalization. The discussion points are detailed in Table 4 and summarized below:

Use the date of service method: HH care is part of the continuum of care; therefore, it should be included in the same episode as the IP hospital stay when measuring physician resource use.

Use the diagnosis method: While many panel members were not in favor of automatically grouping HH claims with the IP stay, there also were significant concerns about the alternative grouping approach based on the HH claims' diagnoses, which many panel members considered to be unreliable and inaccurate.

Exclude HH from episode grouping altogether: Some panel members argued to exclude HH claims from any episodes because too many may have unreliable diagnoses

and may be unrelated to the condition for the hospital stay. Therefore, neither the date of service method nor the diagnosis method should be used to group claims. Also, the variation in HH utilization may be influenced by factors beyond a physician's control, including the degree of local availability of HH services, the degree of social support in the patient's home, and the fact that HH agencies largely control the resource utilization after the initial physician order is given.

Table 4: Feedback on HH Claims

Site 1 (Names blinded)

Home health should be included since it is part of the continuum. Some physicians argued that home health claims should be linked to the hospitalization since it is the triggering event. Their practice is moving towards shifting patients from hospital to home health/hospice. The reason for home health may be unrelated to the reason for the hospital stay. Panel members discussed how it is not possible to know whether the home health stay is connected to the inpatient stay. The patient's needs and resource use will change from the hospital to the home health setting. Home health tends to be more related to the patient's chronic needs than whatever was the original event which led to hospitalization. If the patient was already in home health care prior to a inpatient stay, then the home health episode might be more related to pre-existing conditions and should not be included. Some physicians argued that they had little control over what happens in HH, and therefore HH claims should not be included. Home health claim diagnosis may be inaccurate. Home health diagnoses are not very accurate. The home health diagnosis can often/sometimes reflect the skills (services that) the patient needs, not the true diagnosis. It could be the diagnosis on the home health claim is the one that was needed to get paid for that skill level. Home health agencies have control over resource use in home health, not physicians. The doctor does not determine the specific home health services that are provided to the patient in the home health setting. The home health agency nurse does the assessment and determines what services are needed and how frequent treatments should be. The agency should be held accountable for its efficiency, not the physician. Clinicians responsible for hospital and home health stay may be different. Some physicians argued that responsibility and inclusion in episodes, while related, are different concepts. The home health doctor is not necessarily the same doctor who provided the earlier

treatment in the hospital. After an inpatient stay, home health could be arranged by almost anyone (a discharge planner, hospitalist, and another physician.) For example, an orthopedist may arrange home health after a procedure, but the forms come to the primary care physician to sign off on what was done. A partnership between the two physicians can be expected. In a partnership, a good specialist who orders the home health should also sign off on what was done, but the reality is they have often moved on once the patient has transitioned home.

Home health resource use may be highly influenced by social factors not under the physician's control. It is not fair for physicians to be held accountable for social issues in the home, such as level of family support or the physical layout of the home.

Physician control of home health resource use is not practical currently. The ordering physician should be accountable in the ideal world, but in the real world it is not that practical. This is in the context of a practice plan. It may not necessarily be the same in fee for service Medicare. In an ideal world, the health plan should make sure the cooperation takes place, and not refer to specialists who do not cooperate. Philosophically, this practice would like to see primary care compensated for taking on this level of responsibility. In order to take on that accountability, the delivery model must be reorganized. It would be great if physicians could manage the costs of our patients in the home health setting, but there are not enough hours in the day to manage all that. Today's reality is handoffs with poor coordination and totally fragmented care. Some movements like medical home would be helpful. Episode grouping would line up better with that evolution. Right now, home health is "too far out there" to be included in episodes.

Including or excluding home health claims in episodes could influence physician behavior. To get physicians and PCPs more involved in issues of home health efficiency, including home health in efficiency scores would help motivate it. Again, this prior statement is valid within the context of practice plans. It is much less relevant in a la carte FFS Medicare. Since home health is the least expensive setting in the continuum, excluding it may disadvantage those physicians who rely on it more instead of the more expensive settings. If an episode is determined, in part, by the care setting, then you cannot tell if a physician did a better job in keeping people out of the hospital.

Site 2

Home health is considered part of the continuum of care for hip fracture and COPD patients. A majority of elderly patients will need services in their homes after discharge. Home health is an underutilized resource.

Physicians have formal control over resource use in home health. The doctor is writing the order that determines the specific home health services that are provided to the patient in the home health setting even if the home health agency nurse does the assessment, so the physician should be held accountable for its efficiency.

Determining which physician should be held responsible for home health services is problematic, since in many cases the inpatient physician will be different than the patient's regular physician. There may be widely different approaches from one physician to another, and the physician (frequently a hospitalist) who manages the inpatient care should not be held responsible for the care provided in the outpatient setting (and vice versa). In our practice, these physicians may be separated by great distance and the primary care provider may not even be a physician.

Home health resource use can be highly influenced by social factors not under the physician's control. Some panel members questioned whether it was fair for physicians to be held accountable for social issues, such as level of family support or lack of resources in the community, when they have no control over it. Although this is an issue for all patients, it can be a bigger issue in rural settings.

Site 3

If HH claims are included in episodes, it is preferable to use the primary diagnosis for grouping. The panel members generally were not comfortable using the blunt tool of "date of service" method to link HH to an inpatient stay automatically because Medicare patients are often debilitated for a number of reasons. Rather, if HH is included at all in physician efficiency reporting, then it should be based on the primary diagnosis on the HH claim.

HH should be excluded altogether due to problems with diagnoses on HH claims and physicians' lack of control over HH use. The cardiologists, in particular, voiced the sentiment that the diagnoses often used on HH claims and the physician's very limited control over HH usage were both good arguments for simply excluding HH altogether when measuring episode costs for purposes of physician efficiency reporting.

Site 4

If HH is included in efficiency reporting, use the primary diagnosis to group. Many patients need home health because they are frail, and not because of the inpatient condition. Given that many inpatient hospital claims reflect multiple diagnoses, physicians were also concerned that DRG payment incentives may sometimes affect which diagnosis ends up as the primary diagnosis on the inpatient hospital claim. For both of these reasons, most of the panel members did not think that home health claims should be automatically grouped into the same episode as the inpatient hospital claim.

HH should be included in efficiency measurement, as long as the diagnosis on the HH claim matched what the ordering physician originally coded. Physicians voiced a concern that they have no control over how the HH agency actually submits its claim (e.g., that sometimes the HH claim's primary diagnosis differs from the primary diagnosis the physician specified when ordering the HH care). If the HH agency uses the same diagnostic codes as the physician's primary diagnosis, then it would be appropriate to include HH claims in the same episodes as the inpatient hospitalization.

Related Clinical Events

Identification of Acute Exacerbations

Episodes for a chronic condition can be described as being for maintenance, such as on-going evaluation and management office visits, and/or include acute exacerbations or flare-ups, such as a hospitalization. One episode grouper has an option for separating certain acute exacerbations from the chronic maintenance episodes in five conditions, while the other groups claims for acute exacerbations with chronic maintenance episodes. For example, claims for Acute Myocardial Infarction (AMI) would be grouped into its own acute episode in one episode grouper, but may be included in a coronary arterial disease (CAD) episode in the other. The panels were asked if acute exacerbations should be separated from chronic maintenance episodes, or if they should be included in the same chronic condition episode. The discussion points are detailed in Table 5 and summarized below:

Separate acute exacerbations from chronic maintenance episodes: Acute exacerbations should be broken out from chronic maintenance episodes because the flare-ups are frequently not treated by the physician responsible for a patient's chronic maintenance. Some specialists predominantly treat patients who have an acute exacerbation, and do not treat their ongoing maintenance. An attribution issue could arise for these types of cases if acute exacerbations are not separated. Many factors can

influence the overall management of chronic episodes, such as patient non-compliance, and disease progression, over which a physician has little or no control.

Keep acute exacerbations with the chronic condition: The majority of clinicians favored leaving flare-ups in the same episodes as the chronic condition, because part of a physician's role should be to minimize flare-ups and inpatient admissions. Some clinicians suggested that combining flare-ups with the chronic condition episode may be less of an issue if the risk adjustment methodology is adequate (since sicker patients may be more likely to have acute flare-ups).

Table 5: Feedback on Acute Exacerbations

Site 1 (Names blinded)

Acute exacerbations should be split out because they are frequently not under the physician's control. There is not a perfect correlation between adequate management by the physician and flare-ups. For example, chronic diabetes is all about lifestyle management, and less about how a physician is going to manage the patient. Patient compliance plays a large role in flare-ups of chronic conditions. So this would imply that they should not be included in the episode definition. Acute exacerbations should not be split out because it is part of the physician's role to minimize flare-ups and inpatient admissions. The ultimate outcome is to eliminate the flare-ups and complications. It may be a patient compliance issue, but treating chronic issue is about minimizing the flare-ups. Separating flare-ups from chronic care does not present the full picture of resource use and efficiency. It would be creating negative incentives for any doctor who takes the reporting seriously, because it says he does not have to worry about the flare-ups.

Site 2

Leaving acute exacerbations in with the chronic episode will encourage better patient care. Preventing acute exacerbations is part of the care of chronic conditions. Physicians can make a difference in reducing the number of exacerbations a patient experiences. Including acute exacerbations in with the chronic episode would help to address the perversity in FFS Medicare that poor care receives more reimbursement than good care.

Determining which physician to attribute the episode to would be problematic if the acute exacerbations included in separate episodes. Hospitalists may have no control over what happens to the patient in their chronic care and wants to receive feedback specifically on the care the hospitalist can control. Instead of attributing an episode to only one physician, perhaps it should be attributed to many physicians. Rating individual physicians goes against the team approach of medicine and will not push systems toward efficiency.

Site 3

Splitting out acute exacerbations from the chronic maintenance episodes is important, since many factors influence the management of chronic conditions. The CAD panel emphasized that some cardiologists (including the panel members, who were interventional cardiologists) predominantly treat patients who have already reached an acute exacerbation, and these physicians would need to be evaluated specifically on their care of flare-up cases (and the peer group would also need to include only flare-up cases). The panel members argued that too many other factors influence the overall management of chronic episodes, so the chronic episodes should be separated from the acute flare-ups. For many chronic patients, the hospitalization occurred because the patient either was non-compliant, or simply had a condition that had progressed along an eventual path. In either case the physician had little or no ability to prevent the flare-up.

All inpatient episodes for chronic conditions should be split out as acute exacerbation episodes. For example, although CHF is not one of the conditions for which the grouper currently splits out a flare-up episode, the cardiologists strongly felt that inpatient episodes for CHF should be grouped and evaluated separately from maintenance episodes for CHF.

Site 4

It is not desirable to split out acute exacerbations from the chronic maintenance episodes. If they are split out, physicians may have a financial incentive to avoid (or drop) patients prone to flare-ups. However, if flare-ups are split out then a physician (or group practice) would not get credit for managing their chronic patients to avoid flare-ups. This is part of identifying high-value physicians, one physician noted.

For some conditions it could be difficult to define an acute exacerbation. One physician commented that CHF is already an exacerbation caused by another underlying condition. Under this reasoning, all CHF episodes could be considered an exacerbation. More generally, there was concern about how an acute exacerbation would be defined.

Complications of medical and surgical care

One grouper has an episode type termed "complications of medical and surgical care."[6] In this episode type, claims for a readmission (e.g., a reaction to an implant) will not be grouped into the same episode as the original surgery. The other episode grouper's assignment of the readmission generally will depend on the extent to which that claim's diagnosis appears to match an open episode. Orthopedic surgeons in two panels that focused on hip fractures were asked if claims for complications of surgery should be grouped into the same episode as the original surgery or be grouped into a new episode. The discussion points are detailed in Table 6 and summarized below:

Create a new episode for complications of surgery: If the original surgeon has no control over the readmission, then a new episode should be opened. Some complications are random and due to a patient's co-morbidities and other factors beyond the control of a physician. This arrangement could more easily facilitate attribution, since the same surgeon may not perform both procedures.

Group the complication with the initial episode: Most panelists thought that the complications associated with the original procedure should be included in the same episode as the initial surgery, as doing so could lead to a better measure of physician resource use. By separating the "problem" readmission from the initial procedure, the accountability for the readmission would be lost. One clinician wondered whether risk adjustment methodologies could account for complications. If so, the case for including claims arising from complications with the original surgical episode may be strengthened.

[6] Acumen, LLC found this episode to be usually listed in the top 10 episodes based on aggregate cost.

Table 6: Feedback on Complications of Surgery

Site 1 (Names blinded)

Condition not covered.

Site 2

Including care related to a complication with the original surgery episode depends on the complication. Some complications are connected to the original procedure and should be included. However, if it is a complication that the original surgeon has no control over, it should be split out. If including complications of surgery with the original episode, clean periods may need to be longer. Current periods of 60 to 90 days may not capture complications that could occur up to a year or more. This would be an argument for using at least a one year clean period for a hip surgery. However, grouping all care that occurs within one year with a diagnosis of hip fracture could be problematic. It is unusual to repair two fractured hips at the same time, but not unusual that a second hip would fracture within one year of the first. Thus, if the clean period were extended to a year it would be important to capture whether the surgery was for the same hip or a different hip, and current coding may not capture this information. Splitting complications of surgery out from the original procedure could create negative incentives for providing the best care. For example, if a less expensive prosthesis is used in a hip repair, the physician could look more efficient in the short run, but less efficient in the long run if the prosthesis fails. Sometimes the long run can range from a few years to 10-15 years.

Site 3

If there is a big enough sample size and a good risk adjustment methodology, then complications should be included with the original surgical episode (e.g., with hip fractures). Surgeons were concerned that some complications are simply random, and some are due to a patient's co-morbidities. If both of these concerns can be satisfied, then the surgeons agreed it is best to keep the complications with the original surgical episodes. Early in the discussion the surgeons argued to split out complications, for the reasons above, but as the discussion continued they voiced a stronger concern that relieving accountability for complications could cause some surgeons to opt for less expensive but riskier implants or procedures, with an adverse impact on quality of care. The surgeons thought it could be difficult to agree on the definition of a complication. Several years of data might be required to gain a large enough sample size. Depending on the specific condition being evaluated and the specific surgeon, to get a large enough

sample size it might be necessary to aggregate data across several years, perhaps on a rolling basis.

Site 4

Condition not covered.

Signs and Symptoms (Non-specific) Diagnosis Codes

Signs and symptoms diagnoses are often used to describe presenting conditions, before a final diagnosis of the underlying physiology can be made. Examples of these diagnoses include a cough, and nausea and vomiting. Both groupers attempt to group non-specific diagnosis by using information available from other claims or episodes. If a claim has a non-specific diagnosis code, the panels were asked if the claim should be:

- Grouped with an episode that has a specific diagnosis that is plausibly related and close to the date of service?
- Grouped using a secondary diagnosis (if available)?
- Left ungrouped (would not go to any episode)?

Table 9 details the panel discussions on non-specific diagnosis codes. Clinicians had mixed opinions on how aggressively to group claims with non-specific diagnoses, but they generally favored using only the corresponding diagnosis code to group these claims into episodes. This would leave many of these claims ungrouped (because often there is not an "open episode" that is an obvious choice for grouping the non-specific diagnosis).[7] Table 7 details the discussions on this topic.

[7] Claims with signs and symptom diagnoses that were grouped to the five conditions of interest in this study accounted for approximately 2 percent of the average episode costs for these conditions.

Table 7: Feedback on Signs and Symptoms (Non-specific) Diagnosis

Site 1 (Names blinded)

Sign and symptom coding may reflect the fact that the physician does not know the diagnosis. For example, a chest x-ray may have been part of the process of diagnosing what the problem is. Cough could be coded as the primary diagnosis when the physician does not want to commit to a specific diagnosis. Sign and symptom coding may be influenced by reimbursement issues. The practices' computer system will indicate the importance of using the correct diagnosis for reimbursement purposes. (It is not clear whether this software is following Medicare rules or another insurer's.) Physicians will usually change the diagnosis if it means getting coverage paid for an x-ray. If the x-ray comes back clean, the physician will probably conclude that the patient has COPD. They will put that on their final diagnosis for the visit.

Site 2

It is unclear whether the software should attempt to group claims with non-specific diagnoses. For the purposes of grouping claims, these claims should be included in the episodes. It is unclear if the software is sufficiently sophisticated to determine how and where these codes should be matched, and it is not accurate to group claims based on a presumption. It's difficult to make a decision based on just a few examples. Also, claims information might be insufficient to group correctly. For example, the claims data or the grouping software may not distinguish procedures done on the right hip versus the left hip. Sign and symptom coding may be influenced by reimbursement issues. Non-specific codes may be used in order to be paid, as a more specific but disallowed codes may be rejected.

Site 3

There were mixed views on whether the software should try to group non-specific diagnoses somewhere, or just leave them ungrouped. Several physicians argued that the grouper should not try to "guess" the episode to which these claims should be assigned. One surgeon, however, argued that the grouper should consider all the available information on these claims (including procedures), to try to figure out where to assign these claims. The surgeons indicated that there was very little work done in advance on hip fracture patients.

Site 4

The panel members favored a conservative approach in trying to group claims with non-specific diagnoses. Physicians were uncomfortable with the grouper trying to guess to which episode these non-specific diagnoses should group. They recognized that an alternative policy of automatically leaving all these claims ungrouped might incentivize

physicians simply to code more claims with non-specific diagnoses in order to avoid additional costs in their episodes. There was not much support for using secondary diagnoses from these claims for grouping, since a physician often might note a secondary diagnosis without it being the reason the physician ordered a lab test, x-ray, or other test. The diabetes panel in particular expressed a preference for logic that relies on the primary diagnosis (rather than secondary diagnoses or information on procedures).

An alternative suggested by one physician would be to leave these claims as ungrouped, while also tracking and reporting the percentage of each physician's claims that are ungrouped. This alternative would avoid the grouper "guessing" on these claims, but CMS would have the visibility to track if the ungrouped percentage of physician claims increased significantly over time, for an individual physician, a group of physicians, or even all physicians collectively. If such a trend was observed, it could review its policy regarding how to group these claims.

Other Issues

Duration of Chronic Episodes

By definition, chronic conditions are ongoing and open-ended. To construct an episode to measure resource use, a practical time convention is needed. Both episode grouping software use a 12-month period, usually a calendar year, to measure the costs of chronic conditions. In general, most clinicians were comfortable with using a 12-month period as the convention. Some clinicians thought that the grouper logic should be refined to automatically open a new chronic condition episode in the succeeding 12 month period if the patient had the chronic condition in the past. The grouping software will not open an episode unless a qualifying claim is encountered in the data. The grouping software can only work on the data sets determined by the user, and do not have a "memory" of the prior period data. Table 8 summarizes the feedback at each site.

Table 8: Feedback on the Duration of Chronic Episodes

Site 1 (Names blinded)

It would be important to adjust the costs of a chronic episode for patients who are receiving care for less than the 12-month period. Chronic condition episode costs should be adjusted for patients who have onset mid-year. There are simple ways to adjust for how many months of the episode is open, such as "first in time" methodology used in drug studies. There could be a rule that excludes patients who receive care for less than a 12 month period or a rolling 12 month rule could be used. After a patient is seen the first time for a chronic condition, the 12 month measurement window could be opened. Patient deductibles sometimes impact timing of the first visit of the year. The costs of diabetes two years after onset may be different from the costs of diabetes in the first year after onset. Attribution issues will need to be resolved. There will be questions about how to attribute episode costs for a chronic condition, as patients' providers change over time. One method would be to attribute to the physician that sees the patient first, and then measure 12 months from the date of the first visit for that patient. A 12-month period is probably adequate for measuring resource use. It is possible that 24, 36, or 60 months would be a better measurement/comparison. It might be good to include analyses of 12-month, 36-month, and 60-month experience for each physician. One issue is that initial care for a chronic condition is more resource intensive than maintenance care. Patients with chronic conditions might deteriorate over time. Medicare patients keep getting sicker over time. Good risk adjustment should help a lot with these issues.

Site 2

A 12 month period may be adequate for measuring resource use, but a two year window might be better. Patients with chronic conditions become sicker over time. Initially a COPD patient may be stable, getting a few visits a year and oxygen. When they deteriorate they will need much more care. Patients with a history of chronic conditions should have those episodes automatically reopen at the start of each calendar year. This would apply to all conditions that are chronic – COPD, diabetes, etc.

Site 3

12 months seems a reasonable period to measure, as long as it is the same for everyone.

Site 4

12 months seems a reasonable period to measure a chronic episode. The physicians seemed more concerned with the data lag between the end of the measurement year and when the efficiency reports would be available to physicians. To be actionable, this lag must not be too long.

Exclusion of Certain Claims Types

The grouper software user defines the claim types that are to be grouped. In the episodes constructed for this study, all Medicare Part A and B claims were used. Pharmacy claims in Part D could not be accessed, and therefore, were not included as part of the episode costs. Several claims types were identified for discussion, because they represented claims for which a physician may or may not be held responsible. Table 9 details the discussion that is summarized below on these types of claims:

Durable Medical Equipment (DME) claims: Acumen, LLC analysis found that a large proportion of the DME claims (almost 50% in one grouper) were ungroupable, as a result of the episode groupers' algorithms for grouping claims or the presence of invalid diagnosis on the claims. While many of the DME claims will not be grouped, most panelists thought that DME should be included when measuring physician resource use, even though the diagnoses on DME claims may not be reliable**.**

Ambulance claims: Almost all clinicians believed that ambulance claims should be excluded from an episode because physicians lack control over ambulance usage and

these costs can vary significantly based on geographic factors.[8] Some clinicians expressed particular concern about including the high cost of an air ambulance in an episode. Some clinicians argued that the episode costs should not include any services provided prior to their first contact with the patient for a health condition, since they have no ability to influence the care or treatment until then.

Hospice claims: Physicians in two panels were asked whether hospice care should be included in measuring physician resource use. Most panel members believed that hospice care should be included based on the philosophy that the full continuum of care should be recognized when measuring resource use.

Table 9: Feedback on DME, Ambulance, and Hospice

Site 1 (Names blinded)

DME	
	Physicians have a role in managing DME costs. DME cannot be provided without a doctor's order. DME costs should be included to encourage efficiency. DME has been uncontrolled and unmanaged. DME is much better managed in the managed care health plan than in FFS. So there is a real opportunity for savings. DME costs prior to the first chronic episode claim of the year should be captured. If the grouper could start chronic episodes at the start of year if a patient had it the previous year, then the clean period issue would go away for DME. A grouper feature that automatically opens an episode this year if the patient had a chronic condition the previous year would allow more DME claims to group.
Ambulance	
	Physicians can't control ambulance costs. Ambulance costs should not be included, because the physician has no control over it.

[8] Across the five conditions we focused on in this study, ambulance claims represent approximately 2 percent of the average episode cost.

Ambulance abuse is not a major problem in Medicare. The ambulance cost should be included with the episode. (Most of the time the ambulance claim is valid.) Ambulance costs should be included to encourage efficiency. The practice believes that everything should be included. If you keep these things in, it will drive physicians to practice better.
Hospice
Hospice has new/different group of providers caring for the patient. There are advantages to including end-of-life care and hospice. If you exclude the last months of life, you do not measure the efficiency of physicians who manage the last months better (like getting patients into hospice at the right time). It is better to include end-of-life claims to measure/credit physicians who manage it well. Reducing end-of-life costs can also result in better quality of care, not just cost savings.

Site 2

DME
DME costs prior to the first chronic episode claim of the year should be captured. Can the groupers automatically open a chronic episode at the start of year if a patient had the same chronic condition in the previous year? If so, DME claims would not have to be based on a preceding physician office visit, but on the knowledge of a continuing chronic condition (e.g., monthly oxygen for a COPD patient, who would have an oxygen claim in January but may not have the first physician visit until later in the year).
Ambulance
Physicians can't control ambulance costs. Transportation costs should not be included, because the physician has no control of them.
Hospice
Not discussed.

Site 3

DME
DME claims should be included in the episodes.
Ambulance
Ambulance claims, and any other claims that occur before the physician being evaluated has first seen the patient, should be excluded. This approach was particularly advocated by the surgeons. For example, in addition to excluding ambulance claims, the surgeons wanted to exclude the ER claims for hip fracture, since that has nothing to do with their efficiency as surgeons once they get the patient. The cardiologists had a similar view, although they focused particularly on excluding ambulance claims.

Hospice
Not discussed.

Site 4

DME
DME claims should be included when measuring episode costs for physician efficiency. This opinion was voiced very strongly in the diabetes panel. DME is an important component of diabetes treatment. One panel member described his recommendation to include DME as the most clear-cut of all the issues presented to the panel. The physicians on the CHF panel also supported including DME claims, but some panel members noted that physicians sometimes do not have time to review all of the DME details they are asked to sign off on (and therefore they often have less control of DME than they are supposed to have).
Ambulance
Ambulance claims should be excluded.
Hospice
Hospice and other end-of-life care should be included, as long as there is adequate severity adjustment.

Rural Frontier Issues

The panels at one site believed that their remote location affects practice patterns, increasing costs of treating Medicare patients. Clinicians were also concerned about how valid comparison group costs, such as those used in defining benchmark costs of efficient episodes, would be if resource use reporting did not consider rural frontier issues. Table 10 summarizes the rural frontier issues that were identified by the panels at the Billings Clinic.

Table 10: Billings Clinic Comments on Rural Frontier Issues

Billings Clinic has a catchment radius of approximately 300 miles. Not only do patients travel there for hospital and outpatient services, but the fact that residents of many outlying areas do not have access to long term care, outpatient or home health services in their home communities affects the way care is delivered.

When traveling to Billings for outpatient care, patients tend to save up all their care needs to be attended to in a single outpatient visit or in a series of outpatient visits scheduled for the same day or over a few days. Billings Clinic facilitates and even encourages this practice. Therefore, a potential grouping option that relies solely on date of service for grouping outpatient costs into episodes could be misleading, since patients may see numerous physicians within a few days for unrelated conditions.

Similarly, when patients are hospitalized at Billings the physicians are more likely to attend to their co-morbid conditions while they are inpatient because the patients have to travel long distances for care. Patients typically see numerous physicians for unrelated conditions during an inpatient stay.

Because it is a referral center, Billings Clinic will see patients who have already been seen in another facility's emergency room or who failed to get needed inpatient or outpatient care at another facility. Physicians note that for this reason, their episodes could appear more expensive. For example, if a patient is seen in an outlying ER, hospitalized and then must be transferred to Billings, there is concern that Billings would pick up the earlier facility's costs as part of the episode.

Lack of services near a patient’s home can result in longer inpatient and LTC stays. For example, a patient who could be discharged to home with continued treatment or rehabilitation in an outpatient setting will remain in an inpatient setting because those LTC or rehab services are not available in the patient’s community. When patients do return home it is often to settings in which the Billings provider may have little ability to influence the care received.

Transportation costs, particularly air ambulance, can add significantly to the costs of care, sometimes eclipsing the costs of actual treatment. The use of air ambulances is increased by the need to transport patients quickly over long distances and the reluctance of small communities to having their only ambulance travel out of the community.

Concern was expressed about the adequacy of a comparison group. Such a group should reflect their unique role as a referral center serving a large rural population that must travel long distances to receive care.

Concerns About Using Episode Grouping in Measuring Physician Performance

Overriding concerns about the use of episode grouping in Medicare evolved in all eight panels. These concerns can be categorized as: 1) how cost targets or benchmarks would be constructed, and 2) system attributes. The format for this section of the report differs from the prior section because there is no issue summary and no detailed table of comments for each practice. Since the consensus at all sites was similar regarding these concerns, they could be easily summarized across sites. Because of the many comments about transparency, the format for this one issue is modified to include a detailed table of those comments. The following summarizes the discussions about the concerns from the more than 80 clinicians about measuring physician performance:

Construction of Cost Targets or Benchmarks

- **Logic validation**: Many clinicians mentioned that individual physicians would not have the time or expertise to understand the details of an episode grouper's clinical logic. Consequently, they believed that CMS would have to validate the grouping logic as it applied to Medicare claims before episode groupers could be used in a value based purchasing system.

- **Risk adjustment**: All panels raised concerns about the adequacy of risk adjustment. Panel members viewed risk adjustment as an essential step in using episode grouping for reporting on physician resource use. If physicians believe that the risk adjustment methodology cannot differentiate sicker, more complex and/or frail patients from healthier patients, they will have incentives to avoid treating sicker patients.

- **Homogeneity of episode costs**: Panel members were concerned that the cost for a health condition can differ so greatly within an episode, even in the same practice. In one panel meeting, several orthopedic surgeons asked how very expensive or very inexpensive procedures could be included in a hip fracture episode. Specifically, an inexpensive procedure that uses screws or pins to secure the hip of a seldom ambulatory frail elderly nursing home resident may be indicated, compared to an expensive femur head implant that may be used for a younger, more active person. Since diagnosis is used for grouping, the type of and cost of the procedure is not considered in the grouping process. The orthopedic surgeons questioned how the difference in procedure costs and complexity could be remedied with risk adjustment because the differences in cost and procedure are not a matter of co-morbid conditions.

- **Adequate sample size**: Concern was expressed that the efficiency scores generated with grouped claims data would have an insufficient sample size for statistically validity. The details and criteria for a sufficient sample size were beyond the scope of this study and depend in part on attribution rules. Preliminary analysis of episode grouping data by Acumen, LLC suggests this issue may be a challenge.

- **Valid peer groups**: A strong consensus emerged that great care must be taken when defining the peer group against which a physician or practice would be compared. Of particular concern among the rural group practices was that they would be compared to practices in urban settings, which face different treatment and care issues.

- **Potential for coding changes**: Because the groupers rely heavily on diagnoses, the panels believed that coding practices could be manipulated over time depending upon how resource use reports were used. They suggested that CMS monitor and assess coding changes periodically to determine the extent of providers' ability to gain financially by how coding.

System Attributes

- **Transparency:** Panel members expressed the need for transparency in claim assignments and in how efficiency scores for physicians would eventually be calculated. In the two episode groupers that were studied, claims for medical care of any kind are processed through the episode grouper software and assigned into any of up to roughly 500 different types of episodes. Episode groupers also allow for overlapping simultaneous episodes. The algorithms used to group claims cannot be displayed in a simple to understand branch and tree diagram, such as the DRG system. Because of the complexity of the grouping software, it can be difficult to comprehend how and why the claims are grouped.

  Two important transparency considerations stated in all panels were: 1) how claims are grouped into episodes, and 2) how efficiency scores would be calculated. While most individual physicians will not attempt to work through these methodological issues, they may rely on other professionals to validate the methodologies underlying resource use reports. For physicians in large group practices, the practice's administrative staff and leadership, or national organizations, would likely take the

lead in validating the episode data and grouping processes for their physicians and raising underlying issues to CMS. Physicians in smaller practices will likely rely on their specialty societies to identify broader issues of concern. Because of the overwhelming concern about this issue, Table 11 shows the comments made by the clinicians on this issue.

Table 11: Feedback on Transparency

Site 1 (Names blinded)

Transparency is extremely important. It is important to understand how the grouper builds the episodes with claims data, and how CMS uses that information to build the efficiency scores. The concepts, logic, and analyses need to be explained in easily followed language that physicians can understand. The individual clinicians will look to their specialty societies to inform them how to react, understand the incentives, etc. While the typical individual physician won't get in the "weeds" to understand it, they will react to it. Physicians will need time to understand and respond to efficiency reports. There's a culture of working with that information that evolves over time. The practice introduces metrics, methodologies, reporting, and how they tie performance and compensation to that reporting over a period of years, so that physicians understand it. Transparency will result in Resource Use Reports having greater impact sooner. In the short term, transparency will cause a lot of noise from the docs as they work through the reports. In the longer term, physicians will respond to the incentives. Transparency will drive the process faster.

Site 2

Transparency is extremely important. Because this is going to be attributed to individual physicians, transparency is very important. Physicians will be very defensive and the reaction could be quite contentious. If it's not transparent, CMS will get push back because physicians will not think it is valid. Physicians will want to know how accurate and valid the tools are. They will want to understand how the data are derived. It will not motivate physicians unless the tool has been validated by comparing the results for different types of physician practices (e.g., rural areas, groups affiliated with referral center hospitals, etc.).

Physicians will want to know how the comparison groups are constructed and that they are really being compared to their peers. This concern is especially true of settings in rural areas.

Providers will want to know how their report is relevant to them. Reports will need to be tailored to different types of physicians. The reports must show them what they need to change to be more efficient.

Site 3

Transparency is extremely important.

Physicians would not need to understand the details of episode grouping, but they would want to know that the data are valid. They'll want to know what was done to validate the data and the methodology. It will be also important to analyze the outliers, to see why they are outliers and if the reasons make sense.

Site 4

Transparency is extremely important.

Organizations will care more than individual physicians about the details of grouper logic. Physicians in large group practices will rely on the practice's administrative staff to understand the details of episode grouping logic. Physicians who are not part of a large group practice will rely on their specialty society for this purpose.

Individual physicians will care more about broader issues, such as how their peer group is defined and how the efficiency reports would be used. Physicians will want to know who they are being compared to, and what CMS will do with the results.

- **Actionable information**: Only with actionable information could physicians understand their relative performance and how they could improve their efficiency. Panel members were concerned that a simple bottom line efficiency score would require the need for such detail, but by itself, would not be informative about what kinds of patient management decisions were driving results. Rather, physicians want to know the reasons why their practice was more or less costly than their peers.

- **Quality performance**: Any system for measuring physician efficiency must also include measures of quality, since lower expenditures may result from an unacceptable level of quality. Including quality measures in the relative performance was viewed by panel members as an important component in mitigating potential perverse incentives that might otherwise arise in response to efficiency measurement. Quality adjustment of episodes is conceivably just as important as risk adjustment of episodes.

Conclusion

This report shows clinicians' reactions to eight episode grouper design issues: (1) grouping physician claims with an inpatient hospital stay, (2) grouping an inpatient stay with SNF claims, (3) grouping an inpatient stay with home health claims, (4) excluding certain claims types, (5) grouping complications of medical and surgical care, (6) grouping acute exacerbations of a chronic condition, (7) grouping signs and symptoms (non-specific) diagnosis codes, and (8) defining the duration of chronic episodes. In addition, concerns were voiced about validating the grouper logic, risk adjustment, homogeneity of episode costs, adequate sample size, the validity of peer groups, transparency, actionable information, quality performance, and rural issues. The panel reactions show the importance of bringing persons with clinical knowledge into the development process. The clinician feedback confirms that additional research is needed. This conclusion is consistent with the findings from the Institute of Medicine (IOM) regarding rewarding provider performance: "numerous challenges must be faced in the development, implementation, and ongoing evaluation of performance measures. Multiple methodological considerations—risk adjustment reflecting patient populations of varying acuity, small sample

sizes at the individual practitioner level, … and attribution of responsibility among multiple providers … have already been identified as high priority areas for further research…" (IOM, 2006).

References

AdvanceMed Corporation and TrustSolutions, LLC. **"**Efficiency Measure Special Study: Report on Phase III: Technical Expert Panel Evaluation of Episode" July 31, 2007.

Insitute of Medicine, (September 21, 2006). *Rewarding Provider Performance: Aligning Incentives in Medicare.* Washington, D.C.: National Academy Press. P. 106. Available at http://www.iom.edu/CMS/3809/19805/37232.aspx.

MaCurdy, T., Kerwin, J., Gibbs, J., Lin, E., Cotterman, C., O'Brien-Strain, M., & Theobald, N. (2008-A). "Evaluating the Functionality of the Symmetry ETG and Medstat MEG software in Forming Episodes of Care Using Medicare Data." Burlingame, CA: Acumen, LLC, August 2008. Available at: http://www.cms.hhs.gov/Reports/downloads/MaCurdy.pdf.

MaCurdy, T., Theobald, N., Kerwin, J., Ueda, K. (2008-B). "Prototype Medicare Resource Utilization Report Based on Episode Groupers." Burlingame, CA: Acumen, LLC, November 2008. Available at: http://www.cms.hhs.gov/reports/downloads/MaCurdy2.pdf

Medicare Payment Advisory Commission. Report to the Congress: Assessing Alternatives to the Sustainable Growth Rate System (March 1, 2007). Washington, DC. Available at: http://www.medpac.gov/documents/Mar07_SGR_mandated_report.pdf

Medicare Payment Advisory Commission. March 2005. *Report to the Congress: Medicare Payment Policy.* Washington, DC. Available at: http://www.medpac.gov/documents/Mar05_EntireReport.pdf

Thomas, F. & Caldis, T. (2007) "Emerging Issues of Pay-for-Performance in Health Care," *Health Care Financing Review*, Fall 2007, Vol. 29, pages 1- 4.