Technical Report

A Comparison of the Explanatory Power of Two Approaches to the Prediction of Post Acute Care Resources Use

Jon Eisenhandler, PhD

Richard Averill, M.S.

James Vertrees, PhD

Anthony Quain, M.A.

James Switalski, B.S.

3M Health Information Systems, Inc.

The research described in this report was funded by the Center for Medicare & Medicaid Services under contract HHSM 500-2010-00064C. The opinions expressed in this report are solely those of the authors and do not represent opinions of the Center for Medicare & Medicaid Services.

Published 2011 by 3M Health Information Systems, Inc.

575 West Murray Boulevard, Salt Lake City, Utah 84123
100 Barnes Road; Wallingford, CT 06492
12215 Plum Orchard Drive; Silver Spring, MD 20904

Table of Contents

Executive Summary

The success of Medicare's Inpatient Prospective Payment System (IPPS) in controlling Medicare's cost while not decreasing quality is widely acknowledged. The extension of IPPS to include post acute care (for example, the care provided during the 90 days following discharge from an acute care hospital) has the potential to result in significant savings for Medicare while simultaneously improving quality through better integration of acute and post acute care services.

In order to extend IPPS to include the post-acute period, a risk adjustment method is needed to recognize the impact of a patient's chronic disease burden on the need for post acute care services. A patient's chronic disease burden can be measured based on a patient's pre-existing conditions at the time of hospitalization. There are two methodological approaches to risk adjustment that can, at least in theory, be used to quantify the financial impact of a patient's chronic disease burden during the post acute care period. One is a categorical clinical model such as Clinical Risk Groups (CRGs), the other is a statistical regression-based model such as Diagnostic Cost Group Hierarchical Clinical Conditions (HCCs). The purpose of this study is to compare CRGs and HCCs in terms of their ability to "explain" the variation in post acute care services following a hospitalization as measured by the standard R^2 statistic.

Both CRGs and HCCs were developed to predict costs for an individual for the coming year. While the post acute window of an episode following a hospitalization is in many ways similar to forecasting costs for an individual for the coming year in that both fundamentally involve estimating the effect of disease burden on future costs, there are differences. The standard CRG and HCC models were used in the analysis with no modifications for optimizing them for predicting post acute care services following a hospitalization.

The analyses database encompassed 167 selected MS-DRGs. For this study, post-hospitalization windows of 15, 30, 45, 60, 75, and 90 days were tested. The post acute care services analyzed were hospital outpatient, physician and other part B, DME, skilled nursing facility, home health, hospice and readmissions. Provider charges and Medicare payments were used to measure post acute care service use and were analyzed separately. Payments are less useful as a measure of explained variation than charges because Medicare's PPS payments are based on MS-DRGs and reflect policy and political decision in the existing payment system. A split sample design was used so that one set of records was used to calibrate and a second set of records was used for evaluation.

The SAS HCC program available on the CMS website produces four sets of HCC weights for each individual reflecting the sum of the HCCs, demographic factors, and their initial reason for enrollment (i.e., ever disabled). For this analysis we use the community score as the study population was non-institutional and full enrollment data on every beneficiary was available. The HCC score was assigned to each hospital episode based on the data for the patient in the preceding year including the diagnoses from the trigger hospitalization.

An aggregated version of Clinical Risk Groups (CRGs) comprised of 19 classes (referred to as ACRG4s) was used to create payment levels within each MS-DRG. The CRG was assigned to each hospital episode based on the data for the patient in the preceding year including the diagnoses from the trigger hospitalization. In order for the CRG and HCC comparisons to be on an equivalent basis, annual total charge/payment weights were developed for each of the 19 ACRG4s. The average charges/payments in the calibration database in the one-year period following the hospital discharge that initiated an episode were computed for each ACRG4 and used to create relative weights for each of the 19 ACRG4s. These relative weights were used to predict post acute care services in the evaluation database.

The main test statistic was R^2. Since readmissions are relatively rare, very expensive, and not well predicted by clinical factors, the inclusion of readmissions greatly decreased R^2. For this reason, results are presented including and excluding readmissions. Finally, individuals may not complete an episode because they die or begin another episode. For this reason, the analysis was performed on individuals who “survived” for 90 days – the longest post acute care window evaluated. This means that the same individuals are included in each window. Slightly more than 1.1 million records were used for these analyses.

The fundamental question underlying the analysis is whether either CRGs or HCC can produce stable enough results for use for profiling or payment of post acute episodes. For each MS-DRGs the HCC score and the CRG relative weight was converted to a predicted charge/payment for each patient on a budget neutral basis. The predicted charge/payment was used to compute the R^2 values for HCCs and CRGs.

In summary, the important conclusions from this research include:

- For charges CRGs have a substantially higher R^2 across all windows
- For charges the R^2 for both CRGs and HCCs increases as the length of the window increases but for payments the R^2 is relatively flat as the length of the window increases
- For both CRGs and HCCs the R^2 drops substantially when readmissions are added to the post acute care bundle
- For payments CRGs have substantially higher R^2 for post acute care bundles composed of hospital outpatient, physician and other part B, DME, and home health. However when skilled nursing facility and hospice are added to the post acute care bundle HCCs have a slightly higher R^2
- The correlation coefficient for the predicted CRG and HCC values are 0.612-0.680 for charges and 0.715-0.769 for payments depending on the episode window.

Unlike CRGs, HCCs uses surrogate variables in addition to clinical variables to measure health status. This should bias the HCC performance upward. Despite the use of three non-clinical variables by HCCs, its performance as measured by R^2 is consistently lower.

Chapter One

Introduction

The successful implementation of the Medicare diagnosis related group (DRG) based inpatient prospective payment system (IPPS) in 1983 demonstrated that bundling inpatient services into an all-inclusive, per case payment amounts could create an effective incentive for hospital efficiency by shifting the financial risk for use of bed days and diagnostic and therapeutic services from Medicare to the hospital.

The pressure to further contain Medicare spending is now severe. Because, unlike some Medicare incentive programs, the IPPS incentive structure has proven effective, it can readily be extended to include broader bundles of service that encompass pre- and post-hospital care. While not a panacea, this extension has the potential to result in significant savings coupled with improved outcomes by facilitating a closer integration of inpatient acute and post acute care.

An example of a larger bundle of services is an "Episode of treatment". This is often been defined as the treatment of an illness or condition from beginning to end (Hornbrook, 1985). This approach to defining episodes is focused on events related to the illness rather than focused on the patient, which requires isolating the pre- and post-hospitalization services that were associated with the reason for hospitalization (e.g., all service related to the treatment of diabetes) as distinct from services related to other problems. While the identification of services related to the reason for hospitalization during the pre- and post-hospitalization period for a relatively healthy individual can be done with a reasonable degree of accuracy (e.g., a pregnancy episode encompassing delivery along with pre- and post-partum care, or a cholecystectomy in an otherwise healthy individual), such episodes of care constitute a small proportion of health care expenditures, especially for Medicare beneficiaries. In fact, only about 10% of Medicare beneficiaries consume 63% of Medicare expenditures, (Kaiser Foundation, 2008) and these high utilizing individuals are characterized by multiple co-morbid conditions.

Because the high-utilizing population is characterized by multiple co-morbid conditions, it can be extremely difficult to accurately attribute the pre-hospitalization and post-hospitalization services to the specific condition that was the reason for hospitalization. For example, consider a patient who has congestive heart failure, diabetes, and renal failure. If this patient is hospitalized for uncontrolled diabetes, there will considerable uncertainty in identifying precisely which post-hospitalization services are related to the diabetes care rather than related to the care of the heart failure or renal failure. A post-hospitalization emergency room visit for increasing edema could be attributed to the heart failure, the renal failure, or to the diabetes. The attribution is further complicated because the diabetes is likely the underlying cause of both the heart and renal failure. Since co-morbid diseases interact and do not behave independently, any attempt to isolate only those services that relate to the illness that was the reason for hospitalization will not be accurate for patients with multiple co-morbid conditions (Hughes, 2004). Thus, the disease focused episode concept fails for the very group where increased efficiency could lead to the largest payoffs for the solvency of the Medicare Trust Fund.

Therefore, in order to include persons with multiple chronic conditions, a simpler and more practical definition of an episode is needed. This practical approach to defining episodes should prove more useful in the context of health care reform including physician profiling and payment as it will include those individuals who are of most concern.

A straightforward and practical approach to including beneficiaries with multiple co-morbid conditions in an episode bundle is to define an episode based on the individual (i.e., person-based rather than illness-based). The episode is initiated by the occurrence of a significant health care event such as a hospitalization or a significant ambulatory service. The episode definition can include services within a predefined window of time before and after the event (3 days before and 30 days after, for example). Post-acute care services following a hospitalization are the focus of this paper, since the post-acute care period generally requires a significant amount services that require care coordination. In this approach there is no attempt to assign a service to a particular illness, completely eliminating the intractable problem of determining which visits and procedures are related to the reason for the preceding hospitalization.. This is a person-based episode that is initiated by a trigger event. There are five components needed to insure the success of a patient-centered episode:

- *Episode Trigger*: The event (e.g., hospitalization, ambulatory surgery) that precipitates the episode. A hospitalization for coronary bypass surgery is an example of such an event. This study will use inpatient hospital admissions in selected Medicare Severity Diagnosis Related Groups (MS-DRGs) as trigger events.
- *Episode Acuity:* The severity of the patient's conditions at the time of the event that triggers the episode. Severity is defined as the acuity of the reason for admission as determined by coexisting conditions, and the resultant complexity of care required. This study will use the severity levels in the MS-DRGs as a measure of acuity during the trigger event.
- *Episode Window*: The number of days pre-hospitalization and post-hospitalization that are encompassed by the episode. For example, Medicare currently uses a three day pre-hospitalization window. The post-hospitalization window is a matter of policy and could range from fifteen to as long as ninety days. The incentives for efficient behavior extend to days encompassed within the window so longer windows provide more substantial financial incentives. However, as time increases, the relationship between subsequent events and the trigger event may become less clear.
- *Episode Service Scope*: The services included in the episode (e.g., physician office visits, skilled nursing facility usage, etc.). Incentives for efficient behavior extend to all included services, so larger bundles create more substantial financial incentives. However, if an included service is relatively rare, not clearly associated with the episode types, or relatively expensive, it could shift too much financial risk to providers.

- *Chronic Disease Burden*: The extent of the patient's co-morbid chronic diseases at the beginning of the episode. Since a patient centered episode extends into the post-acute period, a risk adjustment method is needed to capture the chronic disease burden of the individual. The severity of illness and chronic disease burden can be captured by a person-level risk adjustment method based on the person's pre-existing conditions.

Purpose: There are two methodological approaches for risk adjustment that can, at least in theory, be used to quantify the financial impact of comorbid conditions during the post acute care window of patient-based episodes. One is a categorical clinical model such as Clinical Risk Groups (CRGs), the other is a statistical regression-based model such as Diagnostic Cost Group Hierarchical Clinical Conditions (HCCs). The purpose of this document is to compare these two alternatives in the context of person-centered episode definitions in terms of their ability to "explain" the variation in post acute care services following a hospitalization.

Chapter Two

Categorical Clinical Models Versus Regression-based Models

It is first useful to considering the definition of the two alternatives, a categorical clinical model and a regression based formula in somewhat more detail:

1. A *categorical clinical model* consists of a number of discrete cells driven by clinical rules that are defined by clinicians' judgments informed and modified by utilization data so that they are both clinically similar and able to predict or to define an outcome of interest. Cases are classified into mutually exclusive and exhaustive categories based on the patient's underlying health status and applicable demographic factors (e.g. age/sex/disability status). Categorical clinical models are defined by listing specific combinations of clinical and demographic variables that are used to assign patients to a single unique category. MS-DRGs are an example of a categorical clinical model.
2. A *statistical (regression) based model* uses a set of theoretically independent variables to predict the response of a dependent variable that is the outcome of interest, such as costs or mortality. The relationship between the clinical and demographic factors is defined using one of a group of statistical techniques generally referred to as regression models. When the regression model is estimated using a research data set, the result is a set of coefficients (one for each of the independent variables) that can be used to predict the response of the dependent (outcome) variable in an alternate data set. That is, the output of the statistical model is a formula where the mathematical relationships found in the research data set are imposed upon an alternate data set to estimate the value of the dependent (outcome) variable.

Though either method could be used to estimate the impact of the disease burden of an individual on a dependent variable (e.g., the expected cost for post acute care services), what most differentiates the two models is that a categorical clinical model is superior as a communication tool. In fact, achieving the key objectives of improving both efficiency and quality requires changing provider behavior. Indeed, CMS has emphasized the importance of the communications aspect of a categorical model like DRGs to the success of a heath care reform.

> *"The success of any payment system that is predicated on providing incentives for cost control is almost totally dependent on the effectiveness with which the incentives are communicated….Central to the success of the Medicare inpatient hospital prospective payment system is that DRGs have remained a clinical description of why the patient required hospitalization."*
>
> *Federal Register, May 4, 2001*

Although CMS was explicitly referring to systems for cost control, the general point applies to any system that is intended motivate providers to change behavior.

The advantage of a regression-based approach is that it is relatively inexpensive to develop – defining the independent variables is the bulk of the work. However, a regression formula assigns an expected score to each individual and a score is not an effective communication device. While a categorical clinical model is more difficult to develop initially, in addition to facilitating communication, it also has several other advantages.
It simplifies analyses of different dependent variables, especially different types of resources, either alone or in combination. A regression formula must be re-estimated when the dependent variable changes, while a categorical clinical model is unchanging.
It simplifies analyses of different windows and different service scopes. Again, a regression formula must be re-estimated when the time windows change affecting the dependent variable, a categorical clinical model is again unchanging.

As an introduction to the description of the two models (CRGs and HCCs), it will be useful to compare them in terms of their data requirements and in terms of the way that they handle temporal relationships.

Input Data: Both systems are primarily based on diagnoses. Both derive this diagnostic information from administrative data – hospital and physician bills in particular. Both can also use pharmaceutical information (if available) to augment the above information. Both use age/sex categories, albeit very differently, to augment the diagnosis information from the administrative data. The DCG/HCC formula also uses a variable for persons who became Medicare eligible due to disabilities prior to age 65 along with explicit age and sex categories. . This is a surrogate for the chronicity of the individual's disease burden. CRGs use selected procedures to indicate that the patient had a history of major procedure (e.g., history of a heart transplant) and make use of age for a very limited subset of diagnoses. More complete demographic adjustment analogous to the HCC approach is not part of the clinical model and may be implemented to adjust resource estimates.

Data validation rules: DCG/HCC sets a flag for each diagnosis when it first appears irrespective of any other information. In contrast, as is described in more detail below, CRGs use time in the sense of duration for diagnoses reported on physician claims. Physicians often code "rule out" diagnoses so the logic requires that the diagnosis reappear after a pre-determined time has passed (e.g., 90 days) in order for the diagnosis to be considered established. CRGs also use time in the sense of the order of events. For example, a diagnosis may be deleted following a procedure that is expected to cure the problem (e.g., successful coronary bypass surgery should cure angina). However, should the diagnosis reappear following this procedure, CRGs may increase the severity level assigned to the disease (e.g., angina after bypass surgery indicates a failed bypass). Since regression models use dummy (0/1) variables to indicate the presence of a diagnosis, such temporal relationships are difficult to constructs

A Description of the Two Models

The HCC Model: The core of the HCC model is the logic that assigns each diagnosis a position in a hierarchy. All ICD-9-CM diagnoses are assigned to one of 804 initial "DxGroups" each of which represents a similar medical condition. The DxGroups are then grouped into 189 "condition categories" (CCs). Hierarchies are then imposed on the CCs so that, “the most severe manifestation of a given disease process principally defines its impact on costs, … related conditions should be treated hierarchically, with more severe manifestations of a condition dominating less serious ones.” (Pope) The end result of the application of the hierarchy and the fact that some acute categories are not needed for prediction is that only 101 of the CCs (referred to as HCCs) are used as independent variables in the regression model. The DCG/HCC model is parameterized using standard linear regression techniques using the following independent variables:

1. For the prediction model, 101 HCCs. These are summary diagnosis variables that are created in two steps. First, ICD-9-CM diagnosis codes (> 13,000) are collapsed to 804 diagnosis groups. Then these diagnosis groups are further collapsed to 101 HCCs. If an individual has two or more HCCs, the predicted cost is, with the six exceptions noted in 6 below, simply the sum of weights of the HCCs.
2. Twenty four age/sex categories
3. Base year enrollment weight
4. Modified base year enrollment weight for those individuals who were “originally disabled”. (Note that this non-clinical variable is not used by the MS-DRGs or the CRGs.)
5. An adjustment for the “working aged”. (Not used by the MS-DRGs or the CRGs)
6. An adjustment to nine HCCs for the “originally disabled”. (Not used by the MS-DRGs or the CRGs.)
7. An adjustment for six combinations of HCCs. This, in a limited way, allows for certain conditions where costs do not increase in an additive way. Two examples with large interactive effects are between congestive heart failure (CHF) and chronic obstructive pulmonary disease and between diabetes, CHF, and renal failure.

The dependent variable is future cost. Certain anomalies such as negative coefficients are eliminated from the final model.

The CRG Model: CRG development was funded by Department of Commerce, National Institutes of Standards and Technology under Advanced Technology Program. The purpose was to improve the competitive position of the US through the development of a tool that could facilitate managed care. NIST believed that the then currently available risk adjustment methods were relatively ineffective. The first generation methods used age and sex adjustments. The second generation was regression models. CRGs are a 3rd generation risk adjustment system based on diagnostic and treatment history, which use an individual’s medical history and timing in sophisticated ways including onset, duration, sequencing and resolution.

CRGs assign patients to a single mutually exclusive category that predicts the level of overall expected resource use (inpatient and outpatient) during a given time period. Like DRGs, each CRG is composed of a base CRG that describes the patient’s most significant chronic conditions and a severity of illness level (e.g., a patient with diabetes and congestive heart failure at severity level 3). There are 272 base CRGs which are subdivided into up to six severity of illness levels for a total of 1,080 CRGs. Because, like DRGs, CRGs are a “product with a price” model that separates the underlying clinical categorization from the establishment of the associated price (predicted cost), CRGs can provide a measure of the chronic disease burden of a patient at the beginning of an episode. The classification outputs of CRGs and DRGs can be combined to more precisely characterize an episode of care. MS-DRGs can be used to define the severity of the patient's conditions during the episode trigger hospitalization, and CRGs can be used as the basic unit of payment in order to take the chronic disease burden of the patient into account.

The objectives in developing CRGs were to:

- Develop a clinically meaningful means of measuring the health status of a population for the purpose of predicting future health care expenditures
- Develop a management tool for Managed Care Organizations that can also be used for risk adjusting capitated payments
- Develop a language that links the clinical and financial aspects of care

In addition, CRGs contain four to six explicit severity levels within a given category This distinguishes differences in disease burden due to severity of illness, (e.g., not all asthmatics are grouped in the same category). The logic follows the logical progression of a disease. The CRG assignment process is as follows:

Phase 1: **Categorize diagnoses and procedures**

- All diagnoses are assigned to an MDC (Major Diagnostic Category)
- Within each MDCs diagnoses are assigned to one of 565 EDCs (Episode Diagnostic Categories)
- All procedures are assigned to one of 639 EPCs (Episode Procedure Category)
- Each EDC is categorized as dominant chronic, moderate chronic, minor chronic, chronic manifestation, significant acute or minor acute
- Only one diagnosis from an inpatient admission is needed to establish an EDC
- Two diagnoses from different days are needed to establish an EDC for outpatient visits except for diagnoses for selected conditions and diagnosis codes which are in fact procedures (e.g., history of a heart transplant)
- For inpatient services diagnoses from physician and other professional claims are not used (i.e., only the hospital claim is used).
- Diagnoses from "other" providers (e.g., ambulances, freestanding laboratory, etc.) are not used.
- Some diagnosis codes create multiple EDCs. (e.g., the diabetic neuropathy code creates both the chronic disease EDC for diabetes and the chronic manifestation EDC for diabetic neuropathy EDC).
- Conditionality rules are also applied and affect diagnosis or severity assignment:
  - Persistence and recurrence rules (e.g., hypertension must persist over a period of time to be considered an establish diagnosis)
  - Demographic (e.g., congestive heart failure among children vs. adults)
- The temporal relationship between EDCs and EPCs is used to establish final EDCs
  - EDCs can cause other EDCs to be "ignored"

    - Acquired hemiplegia removes stroke from contributing to the severity of illness rating
- EPCs can cause EDC and EPCs to be “ignored”
    - Angioplasty removes Angina from the severity logic
    - Kidney transplant causes renal dialysis to be removed from the severity logic

***Phase 2:* Identify chronic illnesses and specify their severity of illness**

- Each MDC with a chronic EDC will be assigned a PCD (Primary Chronic Disease)
- Only one PCD can be assigned per MDC. If there is more than one EDC within an MDC, the PCDs will be selected in hierarchical order within the MDC (e.g., dominant chronic EDCs selected before moderate chronic EDCs)
- Some chronic EDCs can not become PCDs if a certain other EDC is present (e.g., skin ulcers can not be a PCD if diabetes is present)
- After a PCD is selected it is assigned a severity of illness level
- The severity level assignment for each PCD is establish by the presence of related conditions (e.g., skin ulcers in a diabetic)

***Phase 3:* Assign the CRG**

- Assignment to one of 272 base CRGs based on the combination of PCDs that are present
- The highest volume diseases or combinations of diseases are assigned a unique base CRG, for example:
    - –Diabetes
    - –Diabetes with CHF
    - –Diabetes with CHF and COPD
- All CRGs are assigned to one of nine hierarchical health statuses
- There are nine statuses ranging from catastrophic to healthy
- Assignment is done from most serious (catastrophic) to least serious (healthy)
- Each base CRG is subdivided into discrete severity subclasses based on the severity levels of the PCDs
- Combinations of base CRGs and severity levels result in a total of 1,080 unique clinical groups

***Phase 4:* Assign Adjacent CRGs (ACRGs)**

- The 1,080 CRGs are consolidated into three tiers of aggregation
- Each successive tier of aggregation has fewer base CRGs. Specifically, the number of categories in successive aggregated levels in the current version of CRGs are 416, 151 and 38 referred to as ACRG1, ACRG2, and ACRG3, respectively. As described below for this project a forth level of CRG aggregation with 19 categories was created (ACRG4)
- Severity levels are maintained within each tier.

- Demographic factors such as age, sex, and disability status can be added to the CRGs as an additional adjustment.

Using CRGs for Defining Episodes: Although the application of CRGs for paying for episodes is very similar to the application of DRGs for paying for inpatient care, there are some important differences. DRGs are assigned based on all the diagnoses and procedures that were present at any time during the hospital stay. Thus, DRGs *explain* concurrent hospital resource use based on the care and disease progression of the patient while they were hospitalized. In an episode system, CRGs *predict* episode resource use at the beginning of the episode based on the patient's prior diagnostic and service profile. Since CRGs predict subsequent resource use, they function like a risk adjustment system for capitated payment, which was the original intent.

Essentially, every combination of trigger event, window and service scope defines a unique type of episode. With a categorical episode unit of payment such as CRGs, this diversity is manageable because the process of establishing the projected episode payment amount is straightforward and simply involves computing the historical average resource use of patients in each CRG for each unique type of episode.

In regression based non-categorical systems like HCCs, this diversity is difficult to manage. The coefficients in the regression equation are conceptually equivalent to the average historical resource use (payment weight) in a CRG model. Thus, in a regression model, the score for each individual is equivalent to the category in a categorical model.

Since the ultimate objective of the application of episodes is to create the incentive for cost control and communicate that incentive in a manner and at a level of detail that promotes an effective management response, the issue of a score – which cannot be expressed in a clinically meaningful way - undermines the ability of non-categorical models to provide an effective basis for defining and paying for episodes.

Chapter Three

Data

The initial file contained information for 1,340,820 Medicare beneficiaries who were continuously enrolled in Medicare from 4/1/2006 though 6/30/2009 or the date of their death if they died subsequent to 7/1/2007 with no evidence of another primary payer during that time. Included beneficiaries had three years plus nine months of exposure; including one year prior and 180 days following any hospitalization used as a trigger event.

The purpose of this study was to determine if the method outlined in Chapter 2 could be used to create reasonable episode definitions. Thus, a representative sample was not needed. The trigger event was limited to hospitalizations for 191 selected MS-DRGs V-27. The initial file only included beneficiaries who had a hospitalization between 7/1/2007 and 12/31/2008 that was paid under one of these DRGs. These MS-DRGs were selected based on volume and the expectation that there would be a reasonably consistent pattern of post acute care resource use. These DRGs are identified in Appendix A. Finally, to control the size of the analysis file, the data were limited to nine somewhat diverse (but not random) states. These states – with counts of Medicare beneficiaries in the analysis file are:

- California - 373,169
- Florida – 351,228
- Virginia - 139,228
- New Jersey – 137,834
- Washington - 91,772
- Minnesota – 71,160
- Kansas – 68,732
- Louisiana – 57,356
- Colorado – 50,341

The data include bills for various types of services. These services and the number of bills for each are:

- Inpatient - 4,174,245
- Outpatient - 24,399,272
- SNF - 991,803
- Home health - 1,556,201
- Hospice - 526,376
- DME - 23,913,432
- Part B - 346,061,471

The analysis began with 4,174,245 inpatient hospital claims. Not all of these claims would trigger episodes in the analysis as summarized in Table 1. First, some inpatient claims were hospital transfers; the transfer claims were joined together to arrive at inpatient continuous events. Second, some inpatient continuous events were classified as readmissions of other inpatient stays (i.e., were within the episode window of another inpatient claim); Third, more than half of these inpatient trigger events were outside the analysis period, i.e. i.e. they did not become trigger events for the purposes of this analysis because they did not have sufficient prior history for a CRG to be assigned or did not have sufficient subsequent history for episode window analysis. Fourth, if the patient died during the hospitalization (rather than during the episode window), they were excluded as trigger events. And fifth, inpatient trigger events that were assigned a MS-DRG that was not among the 191 selected DRGs were excluded from the analysis. After making these five adjustments, the number of inpatient trigger events included as episode trigger events was reduced to 1,143,240:

Claims to Episodes	Hospitalizations
Inpatient claims	4,174,245
Transfers	263,359
Readmissions	623,397
Outside analysis period	1,823,981
Hospital deaths	48,660
Excluded DRGs	271,608
Inpatient episodes in analysis database	1,143,240

Table 1: From Inpatient claims to Inpatient Episodes

These inpatient episodes were each analyzed to determine which claims were allocated to each episode, the CRG/HCC risk level of the individual at the beginning of the episode, and the episode costs. This also resulted in the removal of some cases. First, some episodes were not able to have a CRG/HCC assigned (for data quality reasons). Second, if the trigger event charges or payments were less than $200, these cases were dropped from the analysis as an admission costing less than $200 is not credible. Third, cases were dropped based on their MS-DRG if the MS-DRG had low volume (less than 500 cases). In a few instances, low volume MS-DRGs were combined with clinically similar MS-DRGs to obtain a volume of at least 500 cases. Fourth, some MS-DRGs were excluded because later clinical review questioned their appropriateness for inclusion in this project. This analysis was done independently using payments and once using charges (See Table 2).

Inpatient episodes Charges and Payments	Hospitalizations
Inpatient episodes (Charges)	1,143,240
No CRG/HCC assignment	50
Low hospital charges (< $200)	4,184
Low volume/excluded MS-DRG	9582
Usable episodes for charges	1,129,424
Inpatient episodes (Payments)	1,143,240
No CRG/HCC assignment	50
Low hospital payment (< $200)	4,261
Low volume/excluded MS-DRG	9,582
Usable episodes for payments	1,129,347

Table 2: Charges and Payment Data Edits

The following steps were then used to build the analysis file:

- Readmissions can have a substantial impact on post acute costs. In order to avoid having the post acute care cost dominated by a completely unrelated readmission (a subsequent admission for injuries incurred in a traffic accident), a definition of a plausibly related readmission was developed. Any readmission with an MS-DRG in the same major diagnostic category (MDC) as the MS-DRG of the admission that initiated the episode was considered plausibly related to the admission that initiated the episode and was included in the post acute care cost. The one exception to this rule was a list of 49 MS-DRGs that were always considered plausibly related to any admission that initiated episode (see Appendix B). This list was developed by the project clinical team and contains MS-DRGs that are infections and complications of care that could plausibly be related to the care in the admission that initiated the episode. If an unrelated readmission occurred during an episode, the original episode was truncated and a new episode was initiated.

- Only those episodes where an individual beneficiary completed the entire episode were included in the analysis (beneficiary did not die during the episode and did not have an unrelated readmission occur during the episode). Although a method could be easily developed for converting truncated episodes to full episode charges or payments, an adequate number of complete records was available so there was no need to include truncated records for which total charges or payments had to be imputed. The number of hospital episodes excluded from the analysis due to an incomplete episode varied depending on the length of the episode window.

- A split sample design was used so that one set of records was used to calibrate and a second different set of records was used to evaluate the CRGs. This was done by assigning beneficiaries within each MS-DRG a random number. The fifty percent of beneficiaries with the lowest numbers were assigned to the calibration group and the remainder were assigned to the evaluation group.

The analysis is based on the ability of the CRGs/HCCs to predict the Part B services provided in the hospital and all subsequent post acute services during the episode window i.e., services provided after the individual is discharged from the acute care hospital. The data available to this project included two different methods of defining resources: providers' charges and Medicare payments. The charges submitted by the provider on the claim were used for the charge variable. The payment variable was computed as shown in Table 3.

Each of these potential measures of resource use has advantages and disadvantages. Charges likely reflect with more accuracy the relative costliness of individual services. Medicare payments reflect the cost of the service to the program as well as reflecting the outcome of political processes and therefore may or may not reflect real cost differences across services. Since neither is clearly "correct" for all circumstances, the following analyses were done once using charges as the dependent variable and then using Medicare payments as the dependent variable. Other than the dependent variable, the pairs of analyses are identical.

The charges submitted by the provider on the claim were used for the charge variable. The payment variable was computed as follows:

Provider	Description
Hospital	Amount paid with disproportionate share, indirect medical education, new technology add-on amount, and capital removed *plus* beneficiary coinsurance payment *plus* beneficiary deductible payment
Outpatient	Amount paid *plus* beneficiary coinsurance payment *plus* beneficiary deductible payment
SNF	Amount paid *plus* beneficiary coinsurance payment *plus* beneficiary deductible payment
Other part B	Allowed charge
DME	Allowed charge
Home health	Amount paid
Hospice	Amount paid

Table 3: Determination of the payment variable

The CRG is assigned using the diagnoses and procedures present during the hospitalization plus any diagnosis and procedures that occurred one year prior to the date of hospital discharge. The resources that are included in the post acute care episode are those resources that were delivered during the episode window starting on the day following discharge.

Chapter Four

Comparing HCC and CRG based Episodes

The HCC software was downloaded from the CMS website. The file that was used was CMS_HCC_2011Software.zip. This file includes SAS programs, supporting files, and limited documentation. The SAS HCC program produces four sets of HCC weights for each individual reflecting the sum of the HCCs, demographic factors, and their initial reason for enrollment (i.e., ever disabled). For our analysis we chose the SCORE - COMMUNITY as our population was non-institutional and we had full enrollment data on every beneficiary. We tested the HCC as provided in the SAS software with no modifications. While the post acute window of an episode following a hospitalization is in many ways similar to forecasting costs for an individual for the coming year in that both fundamentally involve estimating the effect of disease burden on future costs, there are differences. If the HCC regression was re-estimated to predict post acute care resources, it is reasonable to assume that the performance of this model would, to some extent, improve. However CRGs were also developed to predict costs for an individual for the coming year. So similarly, the CRG clinical model could be optimized for predicting post acute care costs following a hospitalization. Further, the predictive performance of the CRG model would likely be improved if nonclinical factors such as “ever disabled” were incorporated into CRGs. The standard CRG and HCC models were used in the analysis with no modifications for optimizing them for predicting post acute care services following a hospitalization.

The HCC scores were assigned to each hospital episode based on the data for the patient in the preceding year including the diagnoses from the trigger hospitalization. The HCC scores for individual beneficiaries ranged from 0.120 to 19.295 and were used for the comparison. To convert these scores to dollars, they were made “budget neutral” to charges or to payment for each episode. The budget neutrality adjustment was done independent for each type of episode (i.e., MS-DRG) as follows:

1. Sum HCC scores for each type of episode (i.e., MS-DRG).
2. Sum all charges or payments for each type of episode.
3. Divide the sum of charges or payments by the sum of the HCC scores. This budget neutrality factor converts the HCC score to dollars while maintaining “budget neutrality”.
4. Based on step 3, convert the HCC score to dollars for each episode by multiplying the budget neutrality factor times the HCC score. For each enrollee in an episode, this is the HCC predicted (expected) charges or payments for the episode.

Similarly, the CRG was assigned to each hospital episode based on the data for the patient in the preceding year including the diagnoses from the trigger hospitalization. The ACRGs with highest level of aggregation containing 38 unique CRGs was used. For the purpose of evaluating post acute care costs, the number of unique ACRGs was further reduced. At the highest level of aggregation, the nine CRG statuses are subdivided into up to 6 severity levels depending on the status (e.g., healthy has only one level but patients with multiple significant chronic disease have 6 levels). Because all the beneficiaries required hospitalization implying a minimum level severity of illness, the 38 ACRGs were able to be further consolidated into 19 CRG categories as shown in Table 4. The shaded area shows where the severity level is not applicable for a particular status. The numbers in the cells show how the 38 cells were mapped down to the 19 cells. Since all beneficiaries were hospitalized, there were very few patients at status 4 and below so these patients could all be assigned to a single ACRG category. Similarly, beneficiaries at status 5 were able to be consolidated into two ACRG categories. Hospitalized, patients at status 8 and 9 tended to be extremely ill resulting in relatively few patients at the lower severity levels.

The standard HHC scores are intended to predict total annual expenditures. In order for the CRG and HCC comparisons to be on an equivalent basis, annual total charge/payment weights were developed for each of the 19 ACRG4s. The total charges/payments in the calibration database in the one-year period following the hospital discharge that initiated an episode were computed for each beneficiary. Beneficiaries for whom the log of the annual

CRG Status	Severity Level 1	Severity Level 2	Severity Level 3	Severity Level 4	Severity Level 5	Severity Level 6
1. Healthy	1					
2. History of Significant Acute Disease	1					
3. Single Minor Chronic Disease	1	1				
4. Minor Chronic Disease in Multiple Organ Systems	1	1	1	1		
5. Single Dominant or Moderate Chronic Disease	2	2	2	2	3	3
6. Dominant or Moderate Chronic Disease in Multiple Organ Systems	4	5	6	7	8	8
7. Dominant Chronic Disease in Three or More Organ Systems	9	10	11	12	13	13
8. Dominant and Metastatic Malignancies	14	14	14	15	15	16
9. Catastrophic Conditions	17	17	17	18	18	19

Table 4: Mapping of ACRG3s to 19 ACRG4 categories

total charges/payments within an ACRG4 exceeded 2.5 standard deviations above the mean were excluded and the remaining beneficiaries were used to compute the average annual total charges/payments in each ACRG4. The average annual total charges/payments in each ACRG4 were converted to relative weights by dividing by the mean annual total charges/payments across all ACRG4s. The end result was a set of 19 relative weights for charges and 19 relative weights for payment as shown in Table 5. The ACRG4 relative weights based on charges and payments are remarkably similar with only the higher severity levels in the upper statuses showing substantial differences. The ACRG4s that have been consolidated to form the 19 final ACRG4s are assigned the same relative weight in Table 5.

The ACRG4 estimated charge/payment for an episode for each beneficiary in the evaluation database was based on the payment weight in the 19 ACRG4s computed using the calibration database. The same 19 relative weights were used across all episodes. The charge-based weights were used to predict charges and the payment-based weights were used to predict payments. For each type of episode (MS-DRG), a budget neutrality factor was computed as the ratio of the total charge/payment in evaluation subset divided by the sum of the relative weights of the beneficiaries in the episode. For each beneficiary in an episode the relative weight is multiplied by the budget neutrality factor for the episode to compute the ACRG4 predicted (expected) charges/payments. Thus, for each of the 167 episodes (MS-DRGs) there were 19 separate expected charge/payment levels depending on the ACRG assigned to the beneficiary.

Statistical performance is commonly measured by reduction in variance (R^2). As this reflects the reduction in risk for the provider, this statistic is used for this section. In the context of categorical versus regression formula models, it is not clear that there is a method for data trimming that does not favor one or the other method. Therefore, untrimmed data from the evaluation database was used in the comparison.

For the various post acute care windows, the R^2 can be computed in two ways. For each window all the beneficiaries that completed the window (e.g., did not die during the window) could be used as the subset of beneficiaries included in the R^2 computation. As the length of the episode window is increased, this would mean that the number of beneficiaries included in the R^2 computation would decrease so that the R^2 computation for each window would contain a different number of beneficiaries. Alternatively, since the longest episode window examined is 90 days, only those beneficiaries that completed the full 90 day window could be included in the R^2 computation for all windows. Under this method the number of beneficiaries included in the computation of the R^2 is be the same for all windows. This is referred to as the “full 90” approach. Only beneficiaries that completed the full 90 day window were used in the analysis. There were 425,756 beneficiaries in the evaluation database that met the full 90 criteria.

CRG Status	Severity Level 1	Severity Level 2	Severity Level 3	Severity Level 4	Severity Level 5	Severity Level 6
1. Healthy charges	0.339					
1. Healthy payments	0.356					
2. History of Significant Acute disease charges	.0339					
2. History of Significant Acute disease payments	0.356					
3. Single Minor Chronic Disease charges	0.339	0.339				
3. Single Minor Chronic Disease payments	0.356	0.356				
4. Minor Chronic Disease in Multiple Organ System charges	0.339	0.339	0.339	0.339		
4. Minor Chronic Disease in Multiple Organ System payments	0.356	0.356	0.356	0.356		
5. Single Dominant or Moderate Chronic Disease charges	0.436	0.436	0.436	0.436	0.832	0.832
5. Single Dominant or Moderate Chronic Disease payments	0.464	0.464	0.464	0.464	0.874	0.874
6. Dominant or Moderate Chronic Disease in Multiple Organ Systems charges	0.388	0.514	0.619	0.764	1.005	1.005
6. Dominant or Moderate Chronic Disease in Multiple Organ Systems payments	0.428	0.563	0.671	0.825	1.059	1.059
7. Dominant Chronic Disease in Three or More Organ Systems charges	0.658	0.812	1.027	1.250	1.591	1.591
7. Dominant Chronic Disease in Three or More Organ Systems payments	0.712	0.890	1.111	1.319	1.557	1.557
8. Dominant and Metastatic Malignacies charges	0.967	0.967	0.967	1.268	1.268	1.331
8. Dominant and Metastatic Malignacies paments	0.976	0.976	0.976	1.262	1.262	1.274
9. Catastrophic Conditions charges	1.143	1.143	1.143	3.619	3.619	4.508
9. Catastrophic Conditions payments	1.130	1.130	1.130	2.770	2.770	3.293

Table 5: ACRG4 relative weights based on annual totals for charges and payments

Results

Table 6 contains the R^2 by post acute care window for charges with and without readmissions included. The post acute care costs included are hospital outpatient part B, all other part B, DME, home health, skilled nursing facility and hospice. ACRG4s perform consistently better than HCCs. As the length of the window increases, the R^2 increases. The highest R^2 values are for the ACRG4s without readmissions. For ACRG4s the R^2 increases from 14.8 to 28.1 as the window goes from 15 to 90 days. As a reference point, the R^2 for inpatient facility charges for the subset MS-DRGs included in the analysis is 26.4, which is comparable to the ACRG4s for the longer windows without readmissions. The same statistics for HCCs are 8.9 at 15 days and 15.6 at 90 days. Thus, for charges the R^2 for ACRG4s is roughly 66 percent higher than for HCCs. For both ACRG4s and HCCs including readmissions significantly decreases R^2. Indeed, the highest value for ACRG4s is 12.3% while the highest value for HCCs is 7.6%. Readmissions are relatively rare and expensive which do not appear to be closely correlated with either HCCs or ACRG4s so this result is not surprising.

ACRG4 Charges and Payments	Hospital Oupatient	Plus Other Part B	Plus DME	Plus HH	Plus SNF	Plus Hospice	Plus Readmission
Charges	28.9	29.4	29.8	29.9	28.0	28.1	10.6
Payments	22.2	18.9	19.6	18.6	18.0	18.3	12.6

HCC Charges and Payments	Hospital Oupatient	Plus Other Part B	Plus DME	Plus HH	Plus SNF	Plus Hospice	Plus Readmission
Charges	10.7	12.0	12.3	12.6	15.5	15.9	7.6
Payments	9.9	10.9	11.8	12.3	18.4	19.4	13.7

Turning to payments, Table 7 contains the R^2 for the ACRG4s and HCCs for the same post acute windows. As compared to charges, the R^2 values for payments begin higher for ACRG4s and HCCs "without readmissions" but stay relatively flat as the window length increases. The R^2 for HCC is slightly higher than ACRG4s for all windows except 15 days. As with charges, the highest R^2 values are for ACRG4s and HCCs without readmissions.

ACRG4s and HCCs With and Without Readmissions	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
ACRG4 without Readmissions	14.8	20.0	23.0	25.2	27.0	28.1
HCC without Readmissions	8.9	12.3	13.8	14.9	15.6	15.9
ACRG4 with Readmissions	3.2	5.1	6.9	7.3	9.8	10.6
HCC with Readmissions	2.0	3.7	5.0	5.3	7.0	7.6

Table 6: R^2 for post acute charges for ACRG4 and HCCs by episode window (Full90, untrimmed, evaluation database)

Episode Window	15 days	30 days	45 days	60 days	75 days	90 days
ACRG4 without Readmissions	18.0	18.6	18.3	18.0	18.4	18.3
HCC without Readmissions	17.4	19.1	19.4	19.2	19.4	19.4
ACRG4 with Readmissions	8.6	10.8	11.0	12.0	12.1	12.6
HCC with Readmissions	8.2	11.1	11.9	13.0	13.0	13.7

Table 7: R^2 for post acute payments for ACRG4 and HCCs by episode window (Full90, untrimmed, evaluation database)

However, the pattern for ACRG4s for payments is different with R^2 showing a steady increase as the length of the window increases. With readmissions included, R^2 for HCC is slightly higher than ACRG4s for all windows except 15 days

Table 8 contains the R^2 for charges for ACRG4s and HCCs as alternatives as services are added, moving from hospital outpatient and sequentially adding other part B services, DME, home health, skilled nursing facility, hospice and readmissions. In general this pattern of adding services adds less expensive and common services first followed by more expensive and less common services. Each column in Table 8 represents a different episode window.

ACRG4	Post Acute Resources Included Outpatient	Post Acute Resources Included Part B	Post Acute Resources Included DME	Post Acute Resources Included Home Health	Post Acute Resources Included SNF	Post Acute Resources Included Hospice	Post Acute Resources Included Readmission	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
ACRG4	X							15.4	21.4	23.8	26.6	28.1	28.9
ACRG4	X	X						15.6	21.5	24.1	26.6	28.4	29.4
ACRG4	X	X	X					15.7	21.8	24.4	27.0	28.7	29.8
ACRG4	X	X	X	X				15.3	21.5	24.4	27.0	28.8	29.9
ACRG4	X	X	X	X	X			14.8	19.9	22.8	25.0	26.8	28.0
ACRG4	X	X	X	X	X	X		14.8	20.0	23.0	25.2	27.0	28.1
ACRG4	X	X	X	X	X	X	X	3.2	5.1	6.9	7.3	9.8	10.6

HCCs	Post Acute Resources Included Outpatient	Post Acute Resources Included Part B	Post Acute Resources Included DME	Post Acute Resources Included Home Health	Post Acute Resources Included SNF	Post Acute Resources Included Hospice	Post Acute Resources Included Readmission	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
HCCs	X							5.1	7.4	8.8	9.6	10.4	10.7
HCCs	X	X						5.6	8.1	9.5	10.6	11.6	12.0
HCCs	X	X	X					5.7	8.2	9.8	10.9	11.9	12.3
HCCs	X	X	X	X				5.5	8.3	10.0	11.2	12.1	12.6
HCCs	X	X	X	X	X			8.8	12.1	13.4	14.5	15.2	15.5
HCCs	X	X	X	X	X	X		8.9	12.3	13.8	14.9	15.6	15.9
HCCs	X	X	X	X	X	X	X	2.0	3.7	5.0	5.3	7.0	7.6

Table 8: R^2 for post acute charges for ACRG4 and HCCs by episode window and resources included (Full90, untrimmed, evaluation database)

For ACRG4s, R^2 increases as the length of the window increases for all of the alternative service bundles. R^2 is relatively flat as services are added until home health and SNF are added which causes a slight decreased. When readmissions are added there is a significant decrease in R^2 falling from 28.1 to 10.6 for the 90 day window. This pattern is similar, though the R^2 values are always lower, for shorter windows.

The pattern for HCCs is, in part, similar in that R^2 increases as the window length increases. However R^2 starts much lower (10.7% for the 90 day window). Then, unlike for ACRG4s, it moves up to a high of 15.9% as services up to hospice are added and falls to 7.6% as readmissions are included. This pattern implies that HCCs do relative poorly for predicting hospital based outpatient and other part B costs (ACRG4 R^2 2 to 3 times higher) but do relatively better at predicting SNF costs.

ACRG4	Post Acute Resources Included Outpatient	Post Acute Resources Included Part B	Post Acute Resources Included DME	Post Acute Resources Included Home Health	Post Acute Resources Included SNF	Post Acute Resources Included Hospice	Post Acute Resources Included Readmission	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
ACRG4	X							9.9	13.7	16.4	18.9	20.2	22.2
ACRG4	X	X						8.8	13.0	14.7	16.7	17.8	18.9
ACRG4	X	X	X					8.8	13.3	15.1	17.2	18.4	19.6
ACRG4	X	X	X	X				7.3	11.3	13.7	16.0	17.3	18.6
ACRG4	X	X	X	X	X			17.9	18.5	18.0	17.8	18.2	18.0
ACRG4	X	X	X	X	X	X		18.0	18.6	18.3	18.0	18.4	18.3
ACRG4	X	X	X	X	X	X	X	8.6	10.8	11.0	12.0	12.1	12.6

HCCs	Post Acute Resources Included Outpatient	Post Acute Resources Included Part B	Post Acute Resources Included DME	Post Acute Resources Included Home Health	Post Acute Resources Included SNF	Post Acute Resources Included Hospice	Post Acute Resources Included Readmission	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
HCCs	X							4.3	6.0	7.2	8.3	9.0	9.9
HCCs	X	X						4.5	7.0	8.2	9.3	10.1	10.9
HCCs	X	X	X					4.6	7.4	8.7	10.0	11.0	11.8
HCCs	X	X	X	X				4.1	6.9	8.7	10.3	11.4	12.3
HCCs	X	X	X	X	X			17.1	18.6	18.8	18.5	18.7	18.4
HCCs	X	X	X	X	X	X		17.4	19.1	19.4	19.2	19.4	19.4
HCCs	X	X	X	X	X	X	X	8.2	11.1	11.9	13.0	13.0	13.7

Table 9: R^2 for post acute payments for ACRG4 and HCCs by episode window and resources includes (Full90, untrimmed, evaluation database)

Table 9 contains the R^2 for payments for ACRG4s and HCCs as alternatives as services are added. For ACRG4s longer windows are associated with higher R^2 values for all alternative bundles of service. R^2 slowly declines as services are added to the bundle through the addition of SNF (22.2 to 18.3 at 90 days). Adding readmissions reduces R^2 significantly from 18.3 to 12.6 at 90 days.

In Table 9, HCCs have substantially lower R^2 for bundles through inclusion of home health services. However, there is a significant increase in R^2 once SNF services are added (12.3 to 18.4 at 90 days). Once SNF services are added to the bundle, HCCs have a slightly higher R^2 than ACRG4s. This may reflect the use of a disability eligibility variable by HCCs. Individuals who became eligible for Medicare before age 65 due to disability have easier access to nursing homes. For the HCCs as services are added to the bundle. For payment, the R^2 for ACRG4 decreases as services are added to the post acute care bundle but for HCCs the R^2 increases until readmission is added.

The observed difference in R^2 results shows that the predicted charges/payments for ACRG4s and HCCs differ. Table 10 contains the correlation coefficient for the ACRG4 predicted charge/payment and the HCC predicted charge/payment by episode window. While there is a strong relationship between the predicted charge/payment from the two systems, the correlations also indicate that there are differences. In general the correlation for the predicted charge/payment from the two systems is higher for payments than charges, relative stable across windows and higher without readmissions. For charges with a 15 day window with readmissions the correlation is 0.612 while for payments with a 90 day window without readmissions the correlation increases to 0.744. The higher correlation for payments is expected since the R^2 results for ACRG4 and HCCs were more similar for payments than for charges.

Since CRGs were consolidated from 1080 groups down to the 19 ACRG4 groups, the consolidation could have reduced the predictive power of CRGs. Table 11 contains the R^2 for the four levels of CRG consolidation. While there is some reduction in R^2, the reduction is not substantial. The differentiation of beneficiaries by status and severity level provides most of the explanatory value. This is not surprising since for DRGs the large majority of R^2 is due to the differentiation of patients by MDC and the medical/surgical distinction.

The ACRG4 relative weights used in the analysis were based on annual charges/payments and were uniformly applied to every episode (MS-DRG). Because the calculation of the relative weights for CRGs require only the computation of an average value, it is a relatively simple matter to compute relative weights independently for each episode (i.e., change $R(g)$ to $R(e,g)$ in the R^2 formula for CRGs) and compute the relative weights only on post acute care charges/payments. Table 12 contains the R^2 for charges and payments for the two alternative ways of computing ACRG4 relative weights. Episode specific weights based on post acute care charges/expenditures provides a significant increase in R^2, ranging from 10 to 40 percent.

Charges and Payments With and without readmission	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
Charges with Readmission	0.612	0.637	0.639	0.648	0.647	0.654
Payments without Readmission	0.674	0.680	0.673	0.675	0.676	0.674
Charges with Readmission	0.732	0.732	0.732	0.724	0.719	0.715
Payments without Readmission	0.769	0.767	0.764	0.760	0.751	0.744

Table 10: Correlation coefficient between ACRG4 and HCCs predicted post acute care charges and payments by episode window

Without Readmisions	Charges Episode Window 15 days	Charges Episode Window 30 days	Charges Episode Window 45 days	Charges Episode Window 60 days	Charges Episode Window 75 days	Charges Episode Window 90 days	Payments Episode Window 15 days	Payments Episode Window 30 days	Payments Episode Window 45 days	Payments Episode Window 60 days	Payments Episode Window 75 days	Payments Episode Window 90 days
CRG	15.8	21.2	24.4	26.9	28.8	30.1	18.9	19.8	19.5	19.4	19.7	19.6
ACRG1	15.8	21.2	24.4	26.8	28.8	30.1	18.8	19.6	19.3	19.1	19.5	19.3
ACRG2	15.8	21.3	24.5	26.9	28.9	30.2	18.3	19.0	18.6	18.5	18.9	18.8
ACRG3	15.3	20.5	23.6	25.8	27.7	28.9	18.1	18.7	18.3	18.1	18.5	18.4
ACRG4	14.8	20.0	23.0	25.2	27.0	28.1	18.0	18.6	18.3	18.0	18.4	18.3

With Readmisions	Charges Episode Window 15 days	Charges Episode Window 30 days	Charges Episode Window 45 days	Charges Episode Window 60 days	Charges Episode Window 75 days	Charges Episode Window 90 days	Payments Episode Window 15 days	Payments Episode Window 30 days	Payments Episode Window 45 days	Payments Episode Window 60 days	Payments Episode Window 75 days	Payments Episode Window 90 days
CRG	3.4	5.4	7.3	7.7	10.3	11.3	9.0	11.5	11.8	12.9	12.9	13.5
ACRG1	3.4	5.4	7.2	7.6	10.3	11.2	9.0	11.3	11.6	12.7	12.8	13.3
ACRG2	3.4	5.4	7.2	7.6	10.2	11.2	8.7	11.0	11.2	12.3	12.4	12.9
ACRG3	3.3	5.2	7.0	7.4	9.9	10.8	8.6	10.8	11.0	12.1	12.2	12.7
ACRG4	3.2	5.1	6.9	7.3	9.8	10.6	8.6	10.8	11.0	12.0	12.1	12.6

Table 11: R^2 for post acute charges and payment by episode window by CRG level of aggregation

Charge Weights	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
Annual, uniform without Readmissions	14.8	20.0	23.0	25.2	27.0	28.1
Post acute, episode specific without Readmissions	17.8	23.3	26.5	29.1	31.0	32.3
Annual, uniform with readmissions	3.2	5.1	6.9	7.3	9.8	10.6
Post acute, episode specific with readmissions	4.5	6.5	8.7	8.7	11.5	12.3

Payment Weights	Episode Window 15 days	Episode Window 30 days	Episode Window 45 days	Episode Window 60 days	Episode Window 75 days	Episode Window 90 days
Annual, uniform without Readmissions	18.0	18.6	18.3	18.0	18.4	18.3
Post acute, episode specific without Readmissions	23.3	23.4	22.8	22.3	22.3	21.9
Annual, uniform with readmissions	8.6	10.8	11.0	12.0	12.1	12.6
Post acute, episode specific with readmissions	11.6	14.0	14.0	15.1	14.9	15.4

Table 12: R^2 for post acute care charges and payments for ACRG4 by alternative methods of computing relative weights

Discussion

It is important to remember that, unlike CRGs, HCCs uses surrogate variables in addition to clinical variables to measure health status. This should bias the HCC performance upward. Despite the use of three non-clinical variables by HCCs, its performance as measured by R^2 is gradually lower than the clinical CRG model that does not use non-clinical variables. It is important to note that the HCCs were not developed for predicting post acute care following a hospitalization. A statistical process such as using regression methods to estimate a formula results in an optimal model for the original data and the intended circumstances. For HCCs this was predicting next year's Medicare payments.

However, the CRGs were not explicitly developed for this purpose either. However, as a categorical clinical model it can be applied to a wider array of purposes without the need to reformulate the model. As a categorical clinical model, it is not optimized for any data set. Data are only used to verify clinical hypotheses, and clinical judgment always prevails. This is the reason why, for example, that U.S.-developed DRGs been able to be applied to very different health care systems and have been successful world-wide. And this is the reason that CRG's performance was in general superior to HCCs, even when applied for a very different purpose. The superior performance of the CRGs was most evident for charges. Charges as opposed to payments will better reflect the true variation in the amount of post acute care resource use. This very significant superiority of CRGs for charges is indicative of their potential effectiveness in risk adjusting for a wide range of variables beyond payment.

Because the CRG approach subdivides each type of episode into 19 ACRG4s resulting in 3,173 cells (167x19), it raises the question of whether the number of cells can produce an artificially high R^2. The R^2 that would be achieved by randomly splitting the available observations into groups is given by (Feldman, 1992). This is the number of groups minus one divided by the number of observations minus one. Since there are 425,756 hospitalizations that are used to form post acute episodes, the R^2 that would artificially occur based solely on the number of cells is 0.0075. Thus, the R^2 produced by CRGs is due the explanatory power of CRGs and not the number of cells.

Chapter Five

Conclusions

The fundamental question underlying the analysis is whether either CRGs or HCC can produce stable enough results for use for profiling or payment of post acute episodes. In summary, the conclusions from this research include:

- For charges CRGs have a substantially higher R^2 across all windows
- For charges the R^2 for both CRGs and HCCs increases as the length of the window increases but for payments the R^2 is relatively flat as the length of the window increases
- For both CRGs and HCCs the R^2 drops substantially when readmissions are added to the post acute care bundle
- For payments CRGs have substantially higher R^2 for post acute care bundles composed of hospital outpatient, physician and other part B, DME, and home health. However when skilled nursing facility and hospice are added to the post acute care bundle HCCs have a slightly higher R^2
- The correlation coefficient for the predicted CRG and HCC values are 0.612-0.680 for charges and 0.715-0.769 for payments depending on the episode window indicating.

Unlike CRGs, HCCs uses surrogate variables in addition to clinical variables to measure health status. This should bias the HCC performance upward. Despite the use of three non-clinical variables by HCCs, its performance as measured by R^2 is consistently lower than CRGs.

In terms of the issue of whether HCCs or CRGs can produce stable enough results for use for profiling or payment of post acute episodes, the results can be compared to similar results for MS-DRGs. The R^2 for inpatient facility charges for the subset MS-DRGs included in the analysis is 26.4 percent. For CRGs without readmissions included the R^2 for charges ranges from 14.8 percent to 28.1 for windows 15 to 90 days, respectively. So CRGs can produce for a post acute care bundle R^2 values comparable to MS-DRGs. Since MS-DRGs have proved to have achieved an operationally acceptable level predictive performance and a manageable level of financial risk, the R^2 achieved by CRGs would indicate that CRGs should be able to achieve similar results for a post acute care bundle following a hospitalization. However, adding readmissions to the post acute care bundle drops the R^2 for charges substantially to 3.2 percent to 7.6 for windows 15 to 90 days, respectively, which are substantially below the R^2 for MS-DRGs. Thus, the level of financial risk associated with having hospitals at full risk for the cost of readmissions in a post acute care bundle may produce a level of financial risk that is unacceptable and not stable enough results for use for profiling or payment of post acute episodes.

References

Hornbrook, M.C. et al, (1985 Fall). “Health care episodes: definition, measurement and use”, *Med Care Rev.*,42(2):163-218.

Kaiser Family Foundation. (2008).”Distribution of total Medicare beneficiaries and spending, 2005, An analysis of the CMS Current Beneficiary Survey”, *Cost & Use File,* 2005.

Hughes, J.S. et al, (2004), “Clinical Risk Groups (CRGs):A Classification System for Risk-Adjusted Capitation-Based Payment and Health Care Management”, *Med Care Rev.* 42(1):81-90.

Pope, Gregory C., et al, “Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment*”, Final Report under Contract* no. 500-95-048, p.3-3.

Feldman, R, “Are We Splitting Hairs over Splitting DRGs?”, *Health Services Research*, Vol. 27, No. 5, December 1992, 613-617.

Appendix A: MS-DRG Included in Analysis

The sample of beneficiaries included in the data consisted of any beneficiary that was hospitalized in one of 191 MS-DRGs. In the final database used for the analysis 12 of the MS-DRGs were eliminated because of low volume across all severity levels in the MS-DRG or because the MS-DRG was defined based on the beneficiary expiring or leaving against medical advice. An additional 12 MS-DRGs had some of the severity levels with low volume within an MS-DRG consolidated in order to increase patient volume. The final analysis database included the 167 MS-DRGs identified in this Appendix.

MS-DRG	Description	Deleted	Merge Into
001	Heart transplant or implant of heart assist system w MCC	low vol	
002	Heart transplant or implant of heart assist system w/o MCC	low vol	
009	Bone marrow transplant	low vol	
025	Craniotomy & endovascular intracranial procedures w MCC		
026	Craniotomy & endovascular intracranial procedures w CC		
027	Craniotomy & endovascular intracranial procedures w/o CC/MCC		
037	Extracranial procedures w MCC		
038	Extracranial procedures w CC		
039	Extracranial procedures w/o CC/MCC		
064	Intracranial hemorrhage or cerebral infarction w MCC		
065	Intracranial hemorrhage or cerebral infarction w CC		
066	Intracranial hemorrhage or cerebral infarction w/o CC/MCC		
067	Nonspecific cva & precerebral occlusion w/o infarct w MCC		068
068	Nonspecific cva & precerebral occlusion w/o infarct w/o MCC		
069	Transient ischemia		
070	Nonspecific cerebrovascular disorders w MCC		
071	Nonspecific cerebrovascular disorders w CC		
072	Nonspecific cerebrovascular disorders w/o CC/MCC		071
082	Traumatic stupor & coma, coma >1 hr w MCC		084
083	Traumatic stupor & coma, coma >1 hr w CC		084
084	Traumatic stupor & coma, coma >1 hr w/o CC/MCC		
085	Traumatic stupor & coma, coma <1hr w MCC		
086	Traumatic stupor & coma, coma <1hr w CC		
087	Traumatic stupor & coma, coma <1hr w/o CC/MCC		
088	Concussion w MCC		089
089	Concussion w CC		
090	Concussion w CC/MCC		089
100	Seizures w MCC		
101	Seizures w/o MCC		
149	Dysequilibrium		
163	Major chest procedures w MCC		
164	Major chest procedures w CC		

MS-DRG	Description	Deleted	Merge Into
165	Major chest procedures w/o CC/MCC		
175	Pulmonary embolism w MCC		
176	Pulmonary embolism w/o MCC		
177	Respiratory infections & inflammations w MCC		
178	Respiratory infections & inflammations w CC		
179	Respiratory infections & inflammations w/o CC/MCC		
189	Pulmonary edema & respiratory failure		
190	Chronic obstructive pulmonary disease w MCC		
191	Chronic obstructive pulmonary disease w CC		
192	Chronic obstructive pulmonary disease w/o CC/MCC		
193	Simple pneumonia & pleurisy w MCC		
194	Simple pneumonia & pleurisy w CC		
195	Simple pneumonia & pleurisy w/o CC/MCC		
202	Bronchitis & asthma w CC/MCC		
203	Bronchitis & asthma w/o CC/MCC		
207	Respiratory system diagnosis w ventilator support 96+ hours		
208	Respiratory system diagnosis w ventilator support <96 hours		
216	Cardiac valve & oth maj cardiothoracic proc w card cath w MCC		
217	Cardiac valve & oth maj cardiothoracic proc w card cath w CC		
218	Cardiac valve & oth maj cardiothoracic proc w card cath w/o CC/MCC		219
219	Cardiac valve & oth maj cardiothoracic proc w/o card cath w MCC		
220	Cardiac valve & oth maj cardiothoracic proc w/o card cath w CC		
221	Cardiac valve & oth maj cardiothoracic proc w/o card cath w/o CC/MCC		
224	Cardiac defib implant w cardiac cath w/o AMI/HF/shock w MCC		225
225	Cardiac defib implant w cardiac cath w/o AMI/HF/shock w/o MCC		
226	Cardiac defibrillator implant w/o cardiac cath w MCC		
227	Cardiac defibrillator implant w/o cardiac cath w/o MCC		
233	Coronary bypass w cardiac cath w MCC		
234	Coronary bypass w cardiac cath w/o MCC		
235	Coronary bypass w/o cardiac cath w MCC		
236	Coronary bypass w/o cardiac cath w/o MCC		
237	Major cardiovasc procedures w MCC or thoracic aortic aneurysm repair		
238	Major cardiovasc procedures w/o MCC		
242	Permanent cardiac pacemaker implant w MCC		
243	Permanent cardiac pacemaker implant w CC		
244	Permanent cardiac pacemaker implant w/o CC/MCC		
246	Perc cardiovasc proc w drug-eluting stent w MCC or 4+ vessels/stents		
247	Perc cardiovasc proc w drug-eluting stent w/o MCC		
248	Perc cardiovasc proc w non-drug-eluting stent w MCC or 4+ ves/stents		
249	Perc cardiovasc proc w non-drug-eluting stent w/o MCC		
250	Perc cardiovasc proc w/o coronary artery stent w MCC		

MS-DRG	Description	Deleted	Merge Into
251	Perc cardiovasc proc w/o coronary artery stent w/o MCC		
252	Other vascular procedures w MCC		
253	Other vascular procedures w CC		
254	Other vascular procedures w/o CC/MCC		
280	Acute myocardial infarction, discharged alive w MCC		
281	Acute myocardial infarction, discharged alive w CC		
282	Acute myocardial infarction, discharged alive w/o CC/MCC		
283	Acute myocardial infarction, expired w MCC	expired	
284	Acute myocardial infarction, expired w CC	expired	
285	Acute myocardial infarction, expired w/o CC/MCC	expired	
286	Circulatory disorders except AMI, w card cath w MCC		
287	Circulatory disorders except AMI, w card cath w/o MCC		
291	Heart failure & shock w MCC		
292	Heart failure & shock w CC		
293	Heart failure & shock w/o CC/MCC		
294	Deep vein thrombophlebitis w CC/MCC	low vol	
295	Deep vein thrombophlebitis w/o CC/MCC	low vol	
299	Peripheral vascular disorders w MCC		
300	Peripheral vascular disorders w CC		
301	Peripheral vascular disorders w/o CC/MCC		
302	Atherosclerosis w MCC		
303	Atherosclerosis w/o MCC		
304	Hypertension w MCC		305
305	Hypertension w/o MCC		
308	Cardiac arrhythmia & conduction disorders w MCC		
309	Cardiac arrhythmia & conduction disorders w CC		
310	Cardiac arrhythmia & conduction disorders w/o CC/MCC		
311	Angina pectoris		
312	Syncope & collapse		
313	Chest pain		
329	Major small & large bowel procedures w MCC		
330	Major small & large bowel procedures w CC		
331	Major small & large bowel procedures w/o CC/MCC		
335	Peritoneal adhesiolysis w MCC		
336	Peritoneal adhesiolysis w CC		
337	Peritoneal adhesiolysis w/o CC/MCC		
377	G.I. hemorrhage w MCC		
378	G.I. hemorrhage w CC		
379	G.I. hemorrhage w/o CC/MCC		
388	G.I. obstruction w MCC		
389	G.I. obstruction w CC		
390	G.I. obstruction w/o CC/MCC		
391	Esophagitis, gastroent & misc digest disorders w MCC		
392	Esophagitis, gastroent & misc digest disorders w/o MCC		
411	Cholecystectomy w c.d.e. w MCC	low vol	
412	Cholecystectomy w c.d.e. w CC	low vol	

MS-DRG	Description	Deleted	Merge Into
413	Cholecystectomy w c.d.e. w/o CC/MCC	low vol	
414	Cholecystectomy except by laparoscope w/o c.d.e. w MCC		
415	Cholecystectomy except by laparoscope w/o c.d.e. w CC		
416	Cholecystectomy except by laparoscope w/o c.d.e. w/o CC/MCC		
417	Laparoscopic cholecystectomy w/o c.d.e. w MCC		
418	Laparoscopic cholecystectomy w/o c.d.e. w CC		
419	Laparoscopic cholecystectomy w/o c.d.e. w/o CC/MCC		
444	Disorders of the biliary tract w MCC		
445	Disorders of the biliary tract w CC		
446	Disorders of the biliary tract w/o CC/MCC		
459	Spinal fusion except cervical w MCC		460
460	Spinal fusion except cervical w/o MCC		
466	Revision of hip or knee replacement w MCC		
467	Revision of hip or knee replacement w CC		
468	Revision of hip or knee replacement w/o CC/MCC		
469	Major joint replacement or reattachment of lower extremity w MCC		
470	Major joint replacement or reattachment of lower extremity w/o MCC		
471	Cervical spinal fusion w MCC		472
472	Cervical spinal fusion w CC		
473	Cervical spinal fusion w/o CC/MCC		
480	Hip & femur procedures except major joint w MCC		
481	Hip & femur procedures except major joint w CC		
482	Hip & femur procedures except major joint w/o CC/MCC		
490	Back & neck proc exc spinal fusion w CC/MCC or disc device/neurostim		
491	Back & neck proc exc spinal fusion w/o CC/MCC		
492	Lower extrem & humer proc except hip,foot,femur w MCC		
493	Lower extrem & humer proc except hip,foot,femur w CC		
494	Lower extrem & humer proc except hip,foot,femur w/o CC/MCC		
535	Fractures of hip & pelvis w MCC		
536	Fractures of hip & pelvis w/o MCC		
551	Medical back problems w MCC		
552	Medical back problems w/o MCC		
562	Fx, sprn, strn & disl except femur, hip, pelvis & thigh w MCC		
563	Fx, sprn, strn & disl except femur, hip, pelvis & thigh w/o MCC		
602	Cellulitis w MCC		
603	Cellulitis w/o MCC		
637	Diabetes w MCC		
638	Diabetes w CC		
639	Diabetes w/o CC/MCC		
668	Transurethral procedures w MCC		
669	Transurethral procedures w CC		
670	Transurethral procedures w/o CC/MCC		
673	Other kidney & urinary tract procedures w MCC		
674	Other kidney & urinary tract procedures w CC		
675	Other kidney & urinary tract procedures w/o CC/MCC		

MS-DRG	Description	Deleted	Merge Into
682	Renal failure w MCC		
683	Renal failure w CC		
684	Renal failure w/o CC/MCC		
689	Kidney & urinary tract infections w MCC		
690	Kidney & urinary tract infections w/o MCC		
713	Transurethral prostatectomy w CC/MCC		
714	Transurethral prostatectomy w/o CC/MCC		
742	Uterine & adnexa proc for non-malignancy w CC/MCC		
743	Uterine & adnexa proc for non-malignancy w/o CC/MCC		
811	Red blood cell disorders w MCC		
812	Red blood cell disorders w/o MCC		
853	Infectious & parasitic diseases w O.R. procedure w MCC		
854	Infectious & parasitic diseases w O.R. procedure w CC		
855	Infectious & parasitic diseases w O.R. procedure w/o CC/MCC		854
862	Postoperative & post-traumatic infections w MCC		
863	Postoperative & post-traumatic infections w/o MCC		
870	Septicemia or severe sepsis w MV 96+ hours		
871	Septicemia or severe sepsis w/o MV 96+ hours w MCC		
872	Septicemia or severe sepsis w/o MV 96+ hours w/o MCC		
885	Psychoses		
894	Alcohol/drug abuse or dependence, left ama	AMA	
895	Alcohol/drug abuse or dependence w rehabilitation therapy		
896	Alcohol/drug abuse or dependence w/o rehabilitation therapy w MCC		
897	Alcohol/drug abuse or dependence w/o rehabilitation therapy w/o MCC		
919	Complications of treatment w MCC		
920	Complications of treatment w CC		
921	Complications of treatment w/o CC/MCC		

Appendix B: Related Readmissions

Any readmission with an MS-DRG in the same major diagnostic category (MDC) as the MS-DRG of the admission that initiated the episode was considered plausibly related to the admission that initiated the episode and was included in the post acute care cost. The one exception to this rule was a list of 49 MS-DRGs listed in this Appendix that were always considered plausibly related to any admission that initiated episode. The MS-DRGs in this Appendix are infections or complications of care that could plausibly be related to the care in the admission that initiated the episode.

MS-DRG	Description
075	Viral meningitis w CC/MCC
076	Viral meningitis w/o CC/MCC
094	Bacterial & tuberculous infections of nervous system w MCC
095	Bacterial & tuberculous infections of nervous system w CC
096	Bacterial & tuberculous infections of nervous system w/o CC/MCC
097	Non-bacterial infect of nervous sys exc viral meningitis w MCC
098	Non-bacterial infect of nervous sys exc viral meningitis w CC
099	Non-bacterial infect of nervous sys exc viral meningitis w/o CC/MCC
121	Acute major eye infections w CC/MCC
122	Acute major eye infections w/o CC/MCC
152	Otitis media & URI w MCC
153	Otitis media & URI w/o MCC
177	Respiratory infections & inflammations w MCC
178	Respiratory infections & inflammations w CC
179	Respiratory infections & inflammations w/o CC/MCC
193	Simple pneumonia & pleurisy w MCC
194	Simple pneumonia & pleurisy w CC
195	Simple pneumonia & pleurisy w/o CC/MCC
535	Fractures of hip & pelvis w MCC
536	Fractures of hip & pelvis w/o MCC
559	Aftercare, musculoskeletal system & connective tissue w MCC
560	Aftercare, musculoskeletal system & connective tissue w CC
561	Aftercare, musculoskeletal system & connective tissue w/o CC/MCC
602	Cellulitis w MCC
603	Cellulitis w/o MCC
853	Infectious & parasitic diseases w O.R. procedure w MCC
854	Infectious & parasitic diseases w O.R. procedure w CC
855	Infectious & parasitic diseases w O.R. procedure w/o CC/MCC
856	Postoperative or post-traumatic infections w O.R. proc w MCC
857	Postoperative or post-traumatic infections w O.R. proc w CC
858	Postoperative or post-traumatic infections w O.R. proc w/o CC/MCC
862	Postoperative & post-traumatic infections w MCC
863	Postoperative & post-traumatic infections w/o MCC
864	Fever
865	Viral illness w MCC

MS-DRG	Description
866	Viral illness w/o MCC
867	Other infectious & parasitic diseases diagnoses w MCC
868	Other infectious & parasitic diseases diagnoses w CC
869	Other infectious & parasitic diseases diagnoses w/o CC/MCC
870	Septicemia or severe sepsis w MV 96+ hours
871	Septicemia or severe sepsis w/o MV 96+ hours w MCC
872	Septicemia or severe sepsis w/o MV 96+ hours w/o MCC
945	Rehabilitation w CC/MCC
946	Rehabilitation w/o CC/MCC
947	Signs & symptoms w MCC
948	Signs & symptoms w/o MCC
949	Aftercare w CC/MCC
950	Aftercare w/o CC/MCC
951	Other factors influencing health status