Thank you for participating in the Rapid Response Team. In preparation for our next meeting, I wanted to distribute the attached memo from Mathematica Policy Research, Inc. The memo is related to risk adjustment and outlines several issues for which we would like your input:

- What risk factors should be used to capture severity of illness?
- Should regression or standardization methods be used?
- Over what time period should patient history be assessed?
- How should variations in the amount of diagnostic information be managed?
- What risk factors other than diagnosis-based factors should be included?

We will review the memo during our meeting to ensure you have an opportunity to clarify your understanding of the issues and to ask questions.

Response deadline: We will need your feedback on these issues by Monday, March 14 at 4:00 pm. Responses may be provided via email to Katie.burns@state.mn.us.
TO: Katie Burns, MDH
FROM: Eric Schone, MPR
DATE: 3/2/2011
SUBJECT: Risk Adjustment Options

For this project we will prepare reports that compare providers by cost and quality measures. Costs in these reports will be risk adjusted. We are risk adjusting in order to make comparisons between providers that account for the variation in costs that can be expected from treating patient panels with conditions of differing severity. Risk adjustment will also include adjustments related to primary payer type and to variations in the types of services offered by providers. The remaining variation would then be due to differences in the cost of different treatment methods used by providers.

There are two main approaches to risk adjustment. We can reweight a provider’s mix of patients (accounting for differences related to age, gender, and health status) to calculate what the provider’s costs would have been if that provider were treating the same mix of patients as other providers (the “reference population”) and compare that to similarly adjusted costs for other providers, or we can compare the provider’s unweighted actual average costs to the estimated cost that the reference population would have had if they were treating the same mix of patients as that provider.

The first approach, called “direct standardization,” is to divide a provider’s patients into similar categories, calculate the average cost within a category of patients, and then give that group of patients the same weight in the provider’s overall cost average as they have in the total reference population. The adjusted cost can be expressed as

$$A_D = \frac{1}{N_R} \sum_{i=1}^{N_R} \frac{U_{ip}}{N_{ip}}$$

where $N_{iR}$ is the number of patients in category $i$ of $C$ in the reference population, $N_R$ is the total number of patients in the reference population, $N_{ip}$ is the number in the provider’s panel in that same category, and $U_{ip}$ is the total cost from patients in category $i$ in that provider’s panel.¹

¹ Woodward, MD Epidemiology: Study Design and Data Analysis, 1999 Chapman and Hall, Ltd.
MEMO TO: Katie Burns  
FROM: Eric Schone  
DATE: 3/2/2011  
PAGE: 2

For example, if age were the only classifying variable used and Medicare beneficiaries over age 90 make up 20 percent of one provider’s practice but 1 percent of the Medicare population, that provider’s average cost of treating patients over age 90 would get a weight of 1 percent, and the provider’s average cost for patients under 90 would get a weight of 99 percent in calculating the provider’s risk adjusted cost.

The second approach, “indirect standardization,” is to calculate the expected cost by computing the average cost within each category of patients across all the providers in a peer group, then multiplying that average by the proportion of an individual provider’s patients in each category and summing across categories. The result is what we would expect the average cost to be if the provider’s patients were treated by an average provider. If we divide the provider’s actual average cost by this expected average cost, we know how that provider’s results differ from the norm. Multiplying that ratio by the reference population average results in a number on the same scale as the actual cost. The adjusted cost can be expressed as

$$A_t = \frac{U_P / N_P}{\sum_t N_t} \left( \frac{N_{1tR} U_{1tR} / N_{1tR}}{U_{1R}} \right)$$

where $U_R$ is the total unadjusted cost in the reference population, $U_{1R}$ is the reference population’s cost in category I, $N_P$ is the total number of patients in the provider’s panel and $U_P$ is the total unadjusted cost of provider P. Thus, in our age-categorization example above, under this option we would divide the provider’s average actual cost per patient by the weighted average of costs in the reference population for those over and under age 90, using the provider’s weight of 20 percent for those over 90, and 80 percent for those under 90.

Any risk adjustment approach basically employs one of these two approaches. Most of the time, adjustment is performed using the second “indirect” approach, and the expected cost is obtained by a regression, which is a statistical technique used to measure the relationship between patients’ characteristics and their cost across the whole reference population. We plan to use that technique.

We will be risk adjusting two cost measures: 1) total costs based on standardized prices, and 2) raw total costs based on actual payments to providers. We will be risk adjusting costs of two provider types. The first provider type is physicians and the second is hospitals. Physicians’ costs will include all costs incurred by their patients over a certain time period, though costs of patients who are attributed to more than one physician will be attributed proportionally to all of the different physicians. Physician reports will present results aggregated to the clinic or medical group level and individual physicians’ costs will not be identifiable. Hospital costs will include only costs for hospital services associated with a hospital stay, but may include the costs of readmissions to other hospitals.
This memo describes proposed methods and options for risk adjusting these cost measures.\textsuperscript{2} Our proposed methods are briefly summarized below. The remainder of the memo will describe them in more detail.

For physicians: We will identify all service records associated with a patient during a designated period. Using the Johns Hopkins Adjusted Clinical Groups (ACG) system, we will classify the patient into Adjusted Diagnosis Groups (ADGs) according to the diagnoses the patient exhibits during that time period. We will estimate a regression using ADGs and other patient characteristics to predict the patients’ costs and validate the model using bootstrap random samples. We will use the ratio of the physician’s actual costs to expected costs to perform the adjustment.

For hospitals: We will identify all diagnoses for which a patient is treated during the admission and for a designated time period preceding it. We will group these diagnoses using the ACG system and estimate a regression predicting hospital costs, using methods similar to the physician model above. The model will use as predictive factors other patient characteristics and ADGs.

**PHYSICIAN METHODS**

Below, we discuss the following methodological choices for adjusting physician costs:

- What risk factors should be used to capture severity of illness?
- Should regression or standardization methods be used?
- Over what time period should patient history be assessed?
- How should variations in the amount of diagnostic information be managed?
- What risk factors other than diagnosis-based factors should be included?

**What risk factors should be used to capture severity of illness? Should regression or standardization methods be used?**

We plan to use the ACG system to adjust total cost of care to account for risk differences among physician practices. ACGs were developed in the mid-1980s and have been in use in Minnesota since the mid-1990s. The software is used by most health plans in Minnesota for

\textsuperscript{2} We will use actual costs. The only transformation of costs will be exclusion of outliers, as described in “Considerations in Performing Outlier Adjustments”, forthcoming.
multiple purposes, including provider profiling, predictive modeling, and rate setting. ACGs are also used by the Minnesota Department of Human Services for health plan rate setting and medical home payment risk stratification. The system was also used for the nationally recognized program called Patient Choice, which was a managed competition program of competing provider accountable care organizations (ACOs) sponsored by the Minnesota Buyers Health Care Action Group, a purchasing coalition of large employers. It is also used to support the Minnesota State Employee Group Health Insurance Plan’s tiered provider network product called the MN Advantage Health Plan. We will risk adjust the total care cost (standardized and actual) using the ACG Version 9.0 software. We will also use the same software for condition-specific cost measures.

The basic risk adjustment model uses age, sex, ICD-9 diagnosis codes, diagnosis dates, and site of care to assign most types of diagnoses from selected sources to approximately 30 ADGs. Each ADG is assigned based on major diagnostic category, the patient’s prognosis and likely resource costs. For example, one category is “Time Limited”, which identifies a number of acute conditions, while another is “Chronic Medical”. These are then divided into categories, such as minor and major for the time-limited conditions, and stable and unstable, for the chronic conditions. Though each diagnosis is uniquely assigned to one ADG, patients with multiple diagnoses may be assigned to multiple ADGs. Thus, a person can have as many ADGs as are mapped from the person’s diagnosis constellation. The ADGs are analogous to CCs in the Hierarchical Condition Categories (HCC) system used by CMS, in that the diagnosis codes individually map to a single diagnostic category but a patient may be assigned to multiple categories. However, there are many more CCs than ADGs, and clinical similarity plays a large role in CC assignment, while ADG assignment is based to a greater degree on the patient’s likely intensity of care and resulting resource costs.

ADGs serve as the building blocks for the approximately 100 ACGs. The ACGs are based on combinations of ADGs interacted with age, sex and the number of ADGs, to form mutually exclusive categories. Each individual is assigned to a single ACG. Collateral to the ACG system is a classification system that groups diagnoses into approximately 300 Expanded Diagnosis Clusters (EDCs). These categories are assigned on the basis of the clinical relationship between diagnoses rather than their resource implications. For example, it includes categories such as “congestive heart failure” or “other respiratory symptoms” within which severity and costs are likely to be highly variable. EDCs are intended to facilitate exploratory analysis by permitting comparisons of adjusted costs and outcomes for patients with related conditions.
MEMO TO: Katie Burns  
FROM: Eric Schone  
DATE: 3/2/2011  
PAGE: 5

We can use either the ADGs or ACGs for risk adjustment with similar predictive performance. The choice of ACGs or ADGs is related to the choice of estimation by regression or standardization, as described below.

Option 1 ACGs/Standardization

Direct standardization can be used if patients are divided into mutually exclusive groups as they are when ACGs are used. Then the provider’s cost can be standardized by multiplying the provider’s cost per person in a given ACG category times the ACG proportions in the reference population (here Minnesota’s) for each ACG and summing across ACGs. The disadvantage of this approach is that providers may not be able to populate all cells – they will be empty or contain estimates based on only a few observations with correspondingly large variance particularly if their patient population is small.

Option 2 ADGs/Regression

If indirect standardization is used, regression and standardization are similar. For indirect standardization, expected costs are estimated over Minnesota’s population as the mean for each ADG category. Then the physician’s proportions for those ADGs are multiplied by those costs to get an expected cost and the ratio of the physician’s actual cost to expected costs can be calculated. This is equivalent to a regression model that predicts reference population costs as a function of mutually exclusive categories. When risk categories that interact, such as ADGs, are used, multivariate regression techniques instead of standardization techniques are appropriate. Regression methods generally result in more powerful risk adjustment because they permit interacting categories.

Option 3 EDCs/Regression

The EDCs are usually not used for risk adjustment directly. They add little explanatory power, while increasing the complexity of risk adjustment modeling. Therefore, we do not propose to use them for risk adjustment. EDCs will be useful in selecting relevant diagnoses for condition-specific subpopulation reporting.

Recommendation

Because we will calculate risk adjusted rates for physician practices of different sizes—some of which may have relatively small patient populations, and because we expect that modeling interacting ADGs will increase model power, we recommend using regression methods with ADGs as explanatory variables.

---

Over what time period should patient history be assessed?

Option 1: Prospective Risk Adjustment

A prospective model uses a patient’s history for some time prior to the period over which costs are measured, so that events occurring during the measurement period do not affect the expected cost. When risk adjustment is performed prospectively, providers are compared on the cost of managing a patient, given what we know about the patient as of the beginning of the study period. Any information gained during the measurement period is ignored. An advantage of this approach is that worsening of the patient’s condition or treatments provided during the study period might be determined by the provider’s performance. Excluding study-period information ensures that the dependent cost variable is not endogenous to the regression estimation. A disadvantage of this method is that little or no information may be available from the patient history. Even if available, data from a prior year may yield a weak measure of a patient’s current health conditions. Thus, it is likely that the risk adjustment will be too weak.

Because this approach requires all patients and providers have the same risk adjustment information, it is important that a consistent time period be available for each patient. If a prospective model were preferred, we would require at least one year of full claims history prior to the measurement period in addition to whatever period is needed to perform the measurement.

Option 2: Concurrent History – less than one year allowed

Concurrent models explain three or more times the variation compared to prospective models at the individual patient level, because more patient information is available. The greater predictive performance is balanced with an increased concern that the patient’s severity of illness, and the availability of information about that severity, are influenced by treatment decisions. More diagnostic information and potentially higher predicted costs result if the patient receives more services. Also, adverse outcomes resulting from poorly delivered care may result in higher costs, but if concurrent risk adjustment is used predicted costs are higher. Thus, this method is likely to adjust away some differences in provider performance.

An advantage of the concurrent model is that partial year enrollees pose less of a problem for accurate risk assessment than for a prospective model, since information about risk factors is obtained from encounters for treatment. We would typically require a minimum of two-three months enrollment history. A cost per month measure of resource would be used so that all beneficiaries’ costs are compared on an equal basis.

Option 3: Concurrent History – one year only

A concurrent model restricted to patients with one year of enrollment would make the amount of diagnostic information available for risk adjustment more uniform. However, it would reduce the number of patients available to include in each profile, and would remove from the measurement patients who move in and out of the enrolled population during the year, including those who leave due to death.
MEMO TO: Katie Burns  
FROM: Eric Schone  
DATE: 3/2/2011  
PAGE: 7

**Recommendation**
Because we think too much risk adjustment is preferable to too little adjustment for peer grouping purposes, we plan to use concurrent adjustment for the first iteration of peer grouping, and will include partial year enrollees. We will test the use of adjustments for partial year enrollees by estimating the impact on average costs of the average enrollment length of patients attributed to the provider. We may consider shifting to prospective adjustment in future years when more claims data are available.

**How should variations in the amount of diagnostic information be managed?**

When pulling together data from multiple payers, it is likely that we will observe differences in the number of ICD-9 codes per record submitted across payers, even within payer type. Payers may submit up to 15 ICD-9 codes per record. If any payers truncate the number of codes at three or less per record, underestimation of risk for that population may occur. We plan to review the cross-payer data to identify the extent of inconsistency.

- **Option 1 Management not necessary**
  If variation is minor, or if we believe it will have little effect, we would simply make use of all available information.

- **Option 2 Restrict number of diagnoses per claim**
  One response, if diagnoses are severely underrepresented for some payers and the results are strongly affected, is to impose a cap on the number of diagnoses per record. However, such a cap might result in discarding a large quantity of information and weakening the risk adjustment models.

- **Option 3 Include number of diagnoses per claim as a risk factor**
  Including the number of diagnoses per payer or per provider, either as a continuous or dummy variable, in the risk adjustment regressions would permit us to use additional diagnosis codes, while mitigating the bias against any payer that underreports diagnoses.

**Recommendation**
We will review severity as a function of the number of diagnoses permitted. We do not anticipate that the limitation will be a problem, and expect to use all available information. However, if a bias is present, we would anticipate including availability of information in the model in some form.

**What risk factors other than diagnosis-based factors should be included?**

Minnesota’s all-payer claims database has limited direct information about de-identified patient sociodemographic characteristics, including five-digit zip code, age, and gender. Information related to payer type is also available, which serves as a proxy for income.
Option 1 None

The ACG/ADGs include all clinically relevant information and risk adjustment will be based on clinical factors only.

Option 2 Patient sociodemographic characteristics

Some nonclinical patient characteristics such as education or socioeconomic status are beyond the control of providers and influence the likelihood of patient compliance with physician directives or the effectiveness and cost of treatment. These factors may also be associated with the area in which the patient resides. A patient’s primary source of health insurance (commercial health plan, Medicare, or a state-funded public program) serves as a reasonable proxy for these sociodemographic characteristics. Peer grouping reports to providers will include results by primary payer type and public reporting will include a primary payer type adjustment as well. We will review the impact of any additional socioeconomic variables supported by available data by incorporating them in predictive models. We expect, though, that much of the variation in sociodemographics will be associated with variation in payer type and that additional sociodemographic adjustment would be unwarranted.

Option 3 Service mix adjustment

Some stakeholders have expressed concerns that providers with certain characteristics, for example, clinics offering both primary and specialty care rather than solely primary care, will be disadvantaged by profiles that are presented without adjustment for such characteristics. They are concerned that multispecialty clinics will have inherently higher costs because they offer a broader array of services, some of which are more expensive than primary care only services. For the physician clinic total care analysis, the effect of differences in the mix of services offered by providers will be addressed by excluding specialty-only clinics from the total care analysis, limiting the total care analysis to clinics with a certain proportion of primary care clinical staff at the clinic, and adjusting the costs attributed to clinics in the total care analysis by a primary care index. MDH will share more information about this primary care index adjustment in the future.

Recommendation

Our approach to risk adjustment will include adjustments related to clinical factors, primary payer type, and service mix. We will also conduct exploratory analysis to identify which sociodemographic factors for which data are available appear to affect costs even after controlling for ACG/ADGs. If the analyses suggest that some sociodemographic factors have a statistically significant and sizeable effect on costs, we may revisit this decision, after consulting with MDH.
HOSPITAL METHODS

The considerations and methods for adjusting hospital costs and physician costs are generally similar. The unit of measurement for hospital costs is all hospital costs associated with an index admission. Those costs may include the cost of a readmission. The same factors make ADGs using regression methods desirable as the basis for our adjustment. We plan to validate our model, using bootstrapping, if time permits. We plan to investigate variations in the amount of diagnostic information available and control for it if appropriate. In addition, we consider the following methodological choices for hospital risk adjustment that are unique to hospitalizations:

- Over what time period should patient history be measured?
- What risk factors other than diagnosis-based factors should be included?

Over what time period should patient history be assessed?

*Option 1 Prospective Risk Adjustment*

Only risk factors present prior to the index hospital admission would be included in assigning ADGs and developing a risk adjustment model. However, information from the index admission would be excluded as possibly reflecting complications of treatment. Like prospective adjustment of physician costs, this approach requires a consistent history period, and is likely to result in weak adjustment.

*Option 2 Concurrent History – including information from index stay*

All information from the initial stay would be included, excluding diagnosis codes from readmissions. Some of these conditions may occur as a result of treatments being received during the index stay. Because sufficient information could be obtained from the index admission, risk adjustment could be performed even in the absence of a consistent history period. Though some patient risk will be endogenous, this approach results in a stronger risk adjustment model than the prospective approach, and the inclusion of more patients.

*Option 3 Concurrent History – up to one year*

The model would be fit over all information available for a patient during the study period, which may include up to a year of experience following the index stay. This approach includes all available information about the patient. However, it may include information about conditions that developed long after the stay being measured.

---

4 Explained in “Readmissions and attribution of total care costs to hospitals” a memo to Katie Burns dated 9/28/2010.
Recommendation

We recommend using concurrent history, including information from the index stay but not afterward.

What risk factors other than diagnosis-based factors should be included?

Minnesota’s all-payer claims database has limited direct information about de-identified patient sociodemographic characteristics, including five-digit zip code, age, and gender. Information related to the payer type for a particular record is also available, which serves as a proxy for income.

Option 1 None

The ACG/ADGs include all clinically relevant information and risk adjustment will be based on clinical factors only.

Option 2 Patient sociodemographic characteristics

Patients’ sociodemographic characteristics are likely associated with unmeasured burden of illness present on admission and with follow-up services available outside the hospital, which are likely to affect the probability of both discharge and readmission, and consequently the cost of treatment. A patient’s primary source of health insurance (commercial health plan, Medicare, or a state-funded public program) serves as a reasonable proxy for these sociodemographic characteristics. Peer grouping reports to providers will include results by primary payer type and public reporting will include a primary payer type adjustment as well. We will review the impact of any additional socioeconomic variables for which we have data by incorporating them in predictive models. We expect, though, that much of the variation in sociodemographics will be associated with variation in payer type and that additional sociodemographic adjustment would be unwarranted.

Option 3 Service mix adjustment

Some stakeholders have expressed concerns that hospitals vary significantly with respect to the services they provide and that hospitals providing a broader array of services will inherently look more expensive. In addition, hospitals with certain specialized facilities, such as burn units, have higher costs for unrelated conditions, because the costs of these facilities are spread across other revenue centers and bills for other conditions. Similarly, teaching hospitals are likely to have higher costs for all conditions because of the cost of providing medical education in addition to patient treatment. For hospital total care peer grouping, the effect of differences in the mix of services offered by providers will be addressed by creating different peer groups for critical access hospitals and other hospitals, elimination of certain services from the analysis such as trauma and transplants, and outlier truncation, discussed in a previous memo.

Recommendation

Our approach to risk adjustment will include adjustments related to clinical factors, primary payer type, and service mix. However, as for the physician group analysis, we will conduct
exploratory analysis to assess which sociodemographic factors supported by available data (if any) have significant effects on costs, when all diagnostic information is included in the model. Factors that have an independent effect on costs will be considered for addition to the model, after discussion with MDH. However, we do not plan to include certain other provider characteristics in the risk adjustment model, because higher costs for routine hospital admissions due to subsidization of specialized services should not be netted out by risk adjustment. For example, if a hospital with a burn unit has higher costs than other hospitals, controlling for patients’ diagnoses and other factors, the peer grouping analysis should report the higher cost, and the extent to which the difference in hospital costs compares to the estimated marginal effect on costs attributable to having a burn unit. Thus, we plan to conduct sensitivity analysis to identify the effect of provider characteristics such as advanced technical capabilities and teaching status. Our reports would describe the results of this analysis and the impact of these provider characteristics on predicted hospital costs.