White Paper

An Overview of the Risk Adjustment Methodology for PROMETHEUS Analytics©

By Andrew Wilson, MPH, MA

PROMETHEUS ANALYTICS

Introduction

PROMETHEUS Analytics© is perhaps the most comprehensive system available for analyzing provider performance, and developing and implementing value-based payment models around episodes of care. A critical feature of any episode grouper is the application of appropriate methods of risk adjustment to accurately and fairly account for individuals' insurance risk in relation to episode costs. The purpose of this paper is to describe the methodology used to risk adjust episode costs in PROMETHEUS Analytics. It begins with a description of the development history of the method and the key principles that guided its development. The next section lays out the methodology itself, describing the key clinical variables used to model covariates and the modeling procedures that produce expected episode costs. The final section concludes with a discussion of several relevant issues and considerations with respect to the method described.

Development of the PROMETHEUS Analytics Risk Adjustment Methodology

The risk adjustment methodology was developed through a previous collaboration between the Health Care Incentives Improvement Institute (HCI³) and researchers from the Schneider Institutes for Health Policy at Brandeis University to develop an episode grouper for the Centers for Medicare and Medicaid Services (CMS). The framework developed for that project greatly informed the approach described below.

Fundamental Principles of the PROMETHEUS Analytics Risk Adjustment

The risk adjustment process within PROMETHEUS Analytics is based upon several guiding principles:

1. Models should be applicable for multiple uses

The PROMETHEUS Analytics risk adjustment has been developed to give users flexibility in performing risk adjustment for two different use cases. One use is to create fair and accurate comparisons of provider performance by appropriately adjusting for differences in patient severity. Another is to develop patient-specific budgets for bundled payment arrangements based on an individual's insurance risk, such as their demographics and comorbidities.

2. Models should be tailored to the specific patterns of resource use within each individual episode or condition

Because drivers of variation in resource use for a procedure like cataract surgery are very different from those of patients with asthma or those suffering from a stroke, PROMETHEUS Analytics creates different risk adjustment models for each episode of care. This way each model captures the episode-specific contribution of individual risk factors (e.g., age, gender, comorbidities, episode severity) to resource use.

3. Consistent with PROMETHEUS Analytics, the models should distinguish between typical care (i.e., appropriate and patient-centered care) and potentially avoidable complications (i.e., unnecessary or avoidable care).

For each episode of care, the models separately risk adjust costs for typical care and care for potentially avoidable complications (PACs). Segmenting costs this way provides several important advantages. First, from a performance measurement and reporting standpoint, it gives users a way to compare physicians and hospitals along two dimensions of episode costs adjusted for differences in provider case-mix, which can reveal deeper insights about the efficiency and quality of care that comparisons of risk-adjusted average episode costs cannot. Second, for the purposes of bundled payments, the models allow budgets to be constructed that create different yet complementary incentives for individual providers: one that fully rewards providers for providing appropriate typical care for their patients and another that puts significant downward pressure on the occurrence of PACs.

4. Models should create incentives that encourage efficiency and appropriate care and avoid the potential for gaming

PROMETHEUS Analytics risk adjustment avoids creating unwanted incentives in two principle ways. First, the models adjust for only warranted sources of cost variation, or variation that is typical and expected based on the clinical comorbidities of the patient or severity of the procedures being performed. Sources of unwarranted variation, specifically complications and measures of utilization, are expressly avoided as risk-adjusters. Adjusting for complications in the models would remove the incentive to reduce their occurrence. Similarly, adjusting for utilization allows providers whose patients frequently use high intensity services, such as inpatient stays, to appear as better performers or receive higher payments relative to providers whose patients use fewer high cost services.

The second way the models create appropriate incentives is that they rely on prospective risk adjustment to model costs, meaning that the models predict future episode costs using the combination of an individual's comorbidities and markers of episode severity that are known up to and including the point the episode begins. Prospective risk adjustment ensures that the predicted costs for an individual reflect what would be expected given their clinical history. This type of modeling approach is distinctly different from concurrent or retrospective risk adjustment models that utilize diagnoses and events occurring during the episode itself to account for variations in episode costs. Such models are undesirable because they have the potential to promote gaming and introduce incentives that are antithetical to efficiency and quality.

5. The models should be specifically tailored to the user's own data

A wealth of research shows that large differences exist across geographies, payers, and populations in terms of their underlying case mix, fee schedules, coding practices, and provider practice patterns and that these differences contribute to variations in resource use. As a result, not only do costs differ widely for any given episode across populations, so do the risk adjustment models that estimate the unique relationships between the risk factors (i.e., comorbidities, etc.) present in a specific population and its resource use. To ensure that a user's risk adjustment models capture the unique contributions of individual risk factors to episode costs within its own population, PROMETHEUS Analytics requires that personalized sets of risk adjustment models be created each time a new data set is run. Although this may prohibit risk adjustment in some cases when sample sizes for some episodes are limited, it obviates the potential for problems that can arise when applying risk adjustment models developed on "large representative populations" to smaller sub-groups of individuals.

Description of PROMETHEUS Analytics Risk Adjustment Methodology

The PROMETHEUS Analytics risk adjustment models predict individualized episode costs using information on individuals' characteristics, comorbidities, and episode severity. The estimates are the result of a series of regression models that combine to produce expected episode costs. These costs are decomposed into costs for typical care and potentially avoidable complications. Each aspect of the modeling procedure is described in detail below.

Selection of Episodes for Models

By default, episodes are excluded from the models if an individual is less than 18 and greater than 64 years of age, or if the episode is not complete, such as when the episode window extends beyond the time period observed in the claims. To avoid the potential for including false positive episodes or episodes with unusually high costs, episodes in the top and bottom percentiles of the cost distribution are also excluded. Within PROMETHEUS Analytics, users have the option of adjusting the exclusion criteria for their own needs.

Description of Costs

While the episode of care serves as the main unit of inference in analyzing costs, for the purposes of modeling, episode costs are split into costs related to typical care and costs for potentially avoidable complications (PACs). These are determined from the claims assignment process within PROMETHEUS Analytics. For procedural and acute episodes, typical costs are further split into typical costs for the index inpatient stay for the event or procedure; and all remaining typical costs (i.e., professional, outpatient facility, radiology, laboratory, ancillary services and pharmacy costs). All costs for chronic conditions are annualized.

The costs of each component serve as the dependent variables in the models. Separate risk adjustment models are created for each cost component and for every episode type.

Description of Covariates

The risk adjustment models include a number of patient-level covariates. As described earlier, the modeling process is careful to account only for sources of warranted variation that exist up to the point the episode triggers, namely individuals' insurance risk markers. The covariates used in the models include:

- **Patient Demographics and Enrollment:** This includes the individuals' age in years and a dichotomous indicator for gender (1=Female, 0=Male). Also included is an indicator (1=Yes, 0=No) identifying if the individual was enrolled in a plan within the previous six months. The purpose of this variable is to account for these patients' lack of claims history, which limits the number of potential comorbidities that can be identified for the individual.

- **Risk factors:** These include a list of pre-existing comorbidities that an individual has or had prior to the start date of the episode. These are identified from the diagnosis codes that are found in each individual's claims history and are universally applied to every episode type. By default, any risk factor identified over the course of a person's claims history regardless of the time it occurred in relation to the episode can be potentially included in the models. However, users have the option of specifying shorter periods (e.g., 12 months, 18 months, etc.) from which to draw relevant risk factors for the models

- **Episode Subtypes:** These are episode-specific markers that distinguish one episode as being more severe than another. They can indicate either specific patient comorbidities that are known to make a condition or procedure more difficult to treat (e.g., obesity), the severity of the illness itself, or the complexity of the procedure. Subtypes are identified from claims that occurred from the start of the episode up to and including the when the episode triggered. While their inclusion verges on creating a potential for gaming, the subtypes do provide a way to account for important "within episode" differences in severity in the models that aren't captured by pre-existing comorbidities. Moreover, it is far more appropriate to include subtypes at the beginning of the episode than those that would occur later in the episode, which should more appropriately be classified as sequelae of the episode instead of a reclassification of the episode to a different severity level.

- **End-of-Life Probability:** Individuals nearing the end of their lives have very different cost trajectories than those who are not, making end of life a potentially important predictor of episode costs. Simply adjusting for whether an individual died during the episode in the models would create the potential for adverse incentives, as would excluding these episodes from the models altogether. Consistent with the idea of adjusting only for individuals' clinical risk at the beginning of the episode, the models make use of a measure to adjust for potential end of life during the episode. To do this, a logistic regression model is fit to predict a person's probability of death at the beginning of the episode based on their demographics and historical risk factors. These probabilities are then inserted as covariates in the main cost models described in the next section.

Importantly, because the likelihood of death in the commercially insured population is low and commercial claims data frequently lack information on a member's date of death, inclusion of this variable in the risk adjustment is optional.

Once all relevant risk factors and subtypes have been identified for an episode, two additional steps are completed prior to their inclusion in the models. First, to minimize the potential for unstable coefficients, any risk factors or subtypes included in fewer than 10 episodes are discarded for modeling. Second, any overlapping risk factors and subtypes for an episode are collapsed into a single variable.

Modeling Expected Episode Costs

As mentioned, separate risk adjustment models are developed to estimate each of the different cost components (i.e., typical costs, PAC costs) in an episode. The modeling process uses a two-part multivariate regression process to estimate episode costs. In the first step, a logistic regression model is fit to estimate the probability of having non-zero costs for the cost component. Next, ordinary least squares regression (OLS) models are then fit to estimate predicted costs using the subset of episodes with non-zero costs. The full expected costs of the episode cost component are calculated as the product of the probability of having non-zero costs and the predicted costs. Stated differently, the expected costs from these models are the individual's predicted costs conditional on their probability of having any costs for the cost component. The purpose of the two-stage models is to avoid the problems that arise in OLS when a sizeable proportion of observations have the same outcome variable, in this case $0 for the episode-cost component. This causes a violation of the normality assumption underlying OLS and may lead to biased results.

Of note, for some episodes and certain cost components, the two-stage model process is not always needed. For example, most procedure and acute episodes will always have an index stay and other typical costs. However, not all of these episodes may have PAC costs. On the other hand, a chronic condition may have neither typical nor PAC cost if an individual incurs no claims during an observation year. For situations where more than 90% of episodes have non-zero costs for a specific component only OLS models are estimated, otherwise the two-stage method is applied.

Because the purpose of the risk adjustment is to obtain well-developed expected episode costs rather than explain the individual contributions of particular covariates, all risk factors and subtypes meeting the minimum cutoff mentioned previously are included in each of the models. No additional actions are taken to select a more parsimonious set of predictors in the models. This choice reflects a desire to explain as much variation in costs as possible, but it does not make it a priority that all covariates in the models be individually significant or even uncorrelated with each other. It allows for fewer potentially artificial constraints around the definitions of severity of an episode condition, and lets each regression model determine for itself which of the factors are more significant for a specific episode. Non-significant covariates in episode cost models cannot overly influence predicted outcomes, nor is much harm realized, if a group of correlated covariates work together to explain variation rather than having the variation explained by a single best factor.

Once all cost components in an episode have been estimated, these are combined to get a total expected cost for the complete episode.

Additional Issues and Considerations

There are several issues for users to consider when applying the risk adjustment models because they can affect the quality of the analysis and conclusions made from the data. A few of the more important concepts are discussed here.

Model and Covariate Selection

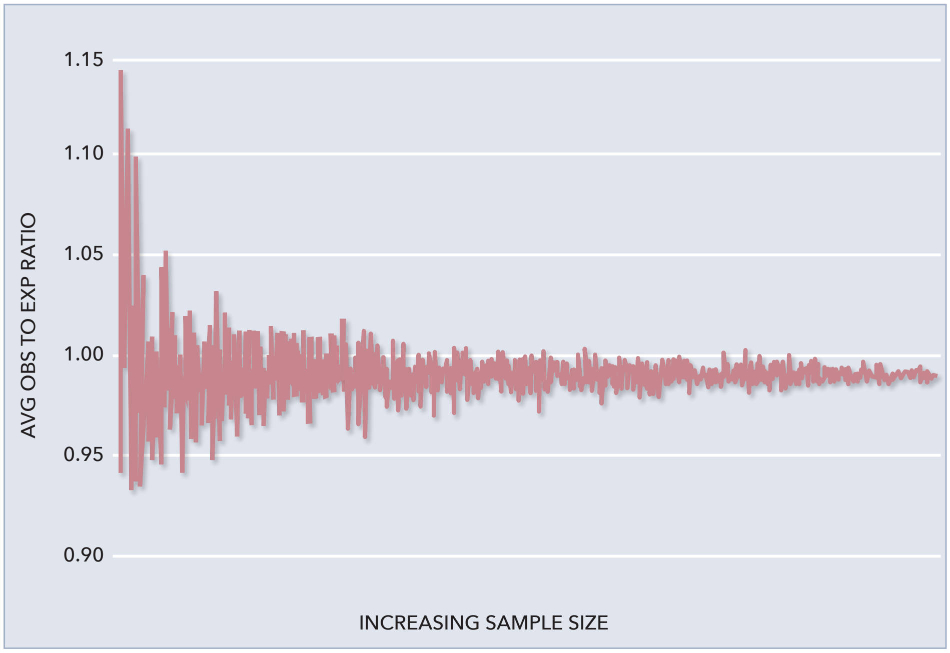
Considerable attention and discussion has been given over the years about how to select the most optimal regression models (e.g., two-part models versus generalized linear models, etc.) as well as selecting the most parsimonious and meaningful set of model covariates (e.g., stepwise regression, regularized regression techniques, etc.) in order to improve cost estimates or minimize the potential for over fitting. There is no consensus on any of these issues and all of the methods available have their own sets of advantages and disadvantages. PROMETHEUS risk adjustment process has sought to strike a balance between statistical rigor and methods that can be implemented and interpreted by a wide range of users. Although more-sophisticated modeling could provide some advantages, such models are not without their own challenges and trade-offs, such as increased complexity and less transparency. Notably, when tested against other modeling approaches in large data sets, the two-part models described above have been found to have almost no differences in terms of the models' predictive power or the potential for overfitting. Nevertheless, the technical processes underlying risk adjustment in PROMETHEUS Analytics can be easily adapted if users are interested in testing and implementing other modeling procedures.

Sample Size

As with all statistical modeling, larger sample sizes produce better estimates of episode costs. Because certain types of episodes occur less frequently than others, especially when data sets are small, it is important for users to consider the number of episodes when applying the risk adjustment models. Low sample sizes can produce models that lack precision in estimating episode costs or lead to over fitting. The issue of poor precision is illustrated in the figure below. It depicts the average prediction error (ratio of average actual costs to average predicted costs) of the risk adjustment model across different sample sizes for knee replacement episodes. At low sample sizes (points closest to the y-axis) the difference between average actual and average expected costs can be as much as 15 percent. As sample size increases, however, the gap between average actual and average expected costs becomes smaller and approaches 1.0, the point where actual and predicted costs are virtually identical.

Practically speaking, what this means is that risk adjustment may not be possible for certain episodes if sample sizes are too small. Unfortunately, there are no specific rules about what an appropriate sample size should be to obtain reasonable models. Rather, such determinations will depend on the population and episodes being studied by users in the context of their data.

FIGURE 1: CHANGE IN PREDICTION ERROR OVER INCREASING SAMPLE SIZE



Relative Provider Prices

Significant variability in pricing between providers and facilities within commercial insurance markets has been extensively documented. While techniques are often used to remove differences in health costs due to pricing variations—usually referred to as "price standardization"—by design these methods are not applied within PROMETHEUS Analytics. This has direct implications on risk adjustment because the expected episode costs predicted by the models will reflect the average price levels within the market. As such, when reporting and comparing risk adjusted episode costs between providers, the performance of providers with high or low prices relative to the market average will be biased even after accounting for differences in case-mix. While PROMETHEUS Analytics views fee-schedule differences as an important source of cost variation to capture and expose, users do have the option to apply a price adjustment factor after severity adjustment to recalibrate the expected cost for high or low priced providers.

Clinical Association of Episodes

One of the key features of PROMETHEUS Analytics is the association of clinically related episodes. For example, costs from the occurrence of a lumbar laminectomy procedure can be attributed back to an individual's underlying case of osteoarthritis. The benefit is that it provides for a complete accounting of all relevant costs incurred within a chronic condition episode for significant procedures as well as for acute medical events. Nevertheless, the addition of these costs, particularly when these costs are substantial, has direct implications on the expected episode costs that are generated by the models.

This issue is illustrated in the table below. The table shows the percentage of osteoarthritis episodes in which a major procedure occurred during the episode and the average increase in costs for osteoarthritis episodes following association of these procedures. The final two columns give the ratio of actual episode costs to the expected episode costs as predicted by the risk adjustment models. From this table, the impact of the clinical associations is clear. The association of the procedures to related osteoarthritis episodes substantially raises actual episode costs. And, as a result, the risk adjustment models under-predict costs for these episodes, sometimes by a great deal. For example, the observed costs of osteoarthritis for individuals who had lumbar laminectomy procedures are just 10 percent higher on average than those predicted prior to the association. But once the procedure is associated this gap increases to about 700 percent.

The point here isn't that the risk adjustment performs poorly in these cases. Indeed, the inability of linear regression to accurately predict extreme values is well understood. Moreover, to the extent that an episode associated to another were otherwise potentially avoidable or unnecessary, the predictions obtained by the risk adjustment models may better reflect what an individual's costs would be had they received the most appropriate care. Rather, the idea is for users to fully understand how to properly implement and interpret the risk adjustment models in the context of one of the fundamental concepts of PROMETHEUS Analytics.

Within the PROMETHEUS Analytics risk adjustment, separate models are created for episodes both prior to and after the association of other episodes.

TABLE 1: IMPACT OF CLINICAL ASSOCIATIONS ON THE COST OF OSTEOARTHRITIS EPISODES

	% OF EPISODES	% INCREASE IN COSTS DUE TO ASSOCIATION	RATIO OF OBSERVED TO EXPECTED COSTS	
			Before Association	After Association
ALL EPISODES	100%	176%	1.0	1.0
EPISODE ASSOCIATIONS				
Lumbar Laminectomy	1%	1,451%	1.1	7.0
Hip Replacement	6%	630%	1.2	1.9
Knee Replacement	9%	571%	1.2	2.0
Knee Arthroscopy	7%	281%	1.1	1.6
Shoulder Replacement	>0%	568%	1.6	4.5

HEALTH CARE INCENTIVES IMPROVEMENT INSTITUTE®
Fair, Evidence-based Solutions. Real and Lasting Change.

13 SUGAR STREET
NEWTOWN, CT 06470
EMAIL: INFO@HCI3.ORG
WWW.HCI3.ORG